MEDICAL LIBRARY

JUL 2 1947

The RESEARCH QUARTERLY

PUBLISHED BY THE AMERICAN ASSOCIATION FOR HEALTH, PHYSICAL EDUCATION, AND RECREATION

MAY, 1947

VOLUME 18
NUMBER 2

The Research Quarterly

of the American Association for Health, Physical Education, and Recreation

MARY WIBEL, *Editor*

MAY, 1947 Volume 18, No. 2

Contents

Page

The Theoretical Efficiency of a Test. Franklin Henry 90

Hidden Possibilities for Research in Physical Education and Recreation. Florence Stumpf and Frederick W. Cozens 104

A Factor-Analysis of Some Cardiovascular-Respiratory Variables and Tests. *Leonard A. Larson* 109

The Relationships of External Fat to Physical Education Activities and *Fitness Tests.* Ramon W. Kireilis and T. K. Cureton 123

The Sexual Maturing of Girls as Related to Growth in Strength. Harold E. Jones 135

The History of School Health Education in West Virginia, 1863-1945. Gladyce H. Bradley 144

Components of Variation and the Consistency of Repeated Measurements. Josef Brozek and Howard Alexander 152

Editorial Policies 167

Reprints from Past Issues of the Quarterly 169

Life and Honorary Members 170

State, Section, and National Officers 173

BOARD OF ASSOCIATE EDITORS

Dorothy Ainsworth
V. S. Blanchard
John Bovard
D. K. Brace
C. L. Brownell
Frederick Cozens
A. Lester Crapser
T. K. Cureton
Leslie W. Irwin
Peter V. Karpovich
C. V. Langton
William K. LaPorte
Leonard A. Larson
Frank S. Lloyd
C. H. McCloy
C. M. Miles
Gertrude E. Moulton
Jay B. Nash
D. Oberteuffer
W. W. Patty
Harry A. Scott
J. R. Sharman
G. T. Stafford
A. H Steinhaus
Alden W. Thompson
W. W. Tuttle
Jesse F. Williams

Published in March, May, October, and December by the American Association for Health, Physical Education, and Recreation, 1201 Sixteenth Street, N.W., Washington 6, D. C. Subscription, $3.00 per year; single copies, $1.00.

Send subscription to: 1201 Sixteenth Street, N.W., Washington 6, D. C.

Editorial office: 1201 Sixteenth Street, N.W., Washington, D. C.

Entered as second-class matter at the Post Office at Washington, D. C., under the act of March 3, 1879.

The Theoretical Efficiency of A Test

By FRANKLIN HENRY
University of California
Berkeley, California

(Submitted for publication, December, 1946)

IT is commonly believed that the effectiveness of a test is proportional to the validity coefficient r (or some index derived from r) relating test scores and criterion scores. Taylor and Russell[7]* have objected to this concept and have in fact shown that a test with a validity coefficient of r = .50 will in some circumstances be more effective than a test with a validity of approximately r = 1.00. Johnson[5] has pointed out that in the case of a 2x2 contingency table analysis, it may be more satisfactory to change the critical passing score than to attempt to increase the validity of a test.

Taylor and Russell[7] have proposed the use of *selection ratio* tables, which make it possible to predict the proportion of success and failure in a test-selected group of individuals if the following information is available:

1. The proportion of unselected individuals who are successful in the criterion situation.
2. The coefficient of correlation relating unselected test scores to criterion success and failure.
3. The proportion of individuals who are retained in the selected group.

This method of evaluating test efficiency proved quite valuable in the Army Air Forces' wartime psychological research on aircrew selection. An extension of the method was also found useful in estimating the value of tests of fitness for flying at extremely high altitudes.[1] In the latter case, insight into testing theory gained by the application of this method made it possible to reconcile differences in experimental results that appeared quite conflicting when analyzed by conventional methods.[4]

It is the purpose of the present study to show that if the selection-ratio method is derived on a broader basis than was considered by Taylor and Russell, it is possible to investigate phases of the problem that were outside the scope of their study. Two aspects will be of particular interest, namely:

From the Department of Physical Education and the Division of Medical Physics. The work described in this report was supported, in part, by a contract recommended by the Committee on Medical Research, between the Office of Scientific Research and Development, and the University of California.

*Superior figures refer to numbered bibliography at end of article.

1. Delineation of the relationships between degree of selection and degree of criterion success for specified validity coefficients, leading to a rational basis for determining the best placement of "cutting scores" for selection or rejection of tested individuals, and for deciding whether it is worthwhile to use a certain test in some particular situation.

2. Determination of the effectiveness of a test in eliminating potentially successful individuals from the *rejected* group. This aspect is extremely important in some situations, as for example when individuals are excluded from professional training on the basis of test battery scores, or are advised by a vocational counselor to avoid certain trades or professions.

It should be pointed out that there are situations in which the results of the present study do not apply. In research applications, for example, interest may be centered in the *relation* between test scores and criterion scores. However, these applications are not directly concerned with *selection*. But there are many other situations in which success or

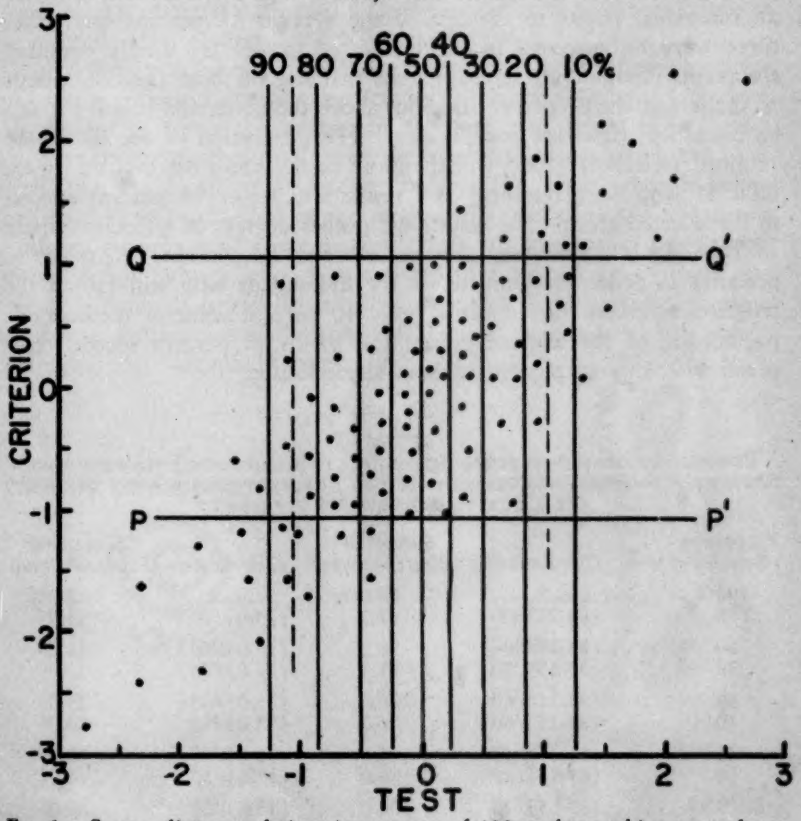

FIG. 1. Scatter diagram relating sigma scores of 100 students taking a one hour test to sigma scores on a criterion consisting of a three-hour comprehensive examination. The validity coefficient is r = .80.

failure must necessarily be defined in terms of pass or fail. An individual does or does not graduate from school; he does or does not win a race; he does or does not get or lose a job, have an accident, or become married or divorced.

SOME FUNDAMENTAL RELATIONSHIPS

In Figure 1 is shown a scatter diagram relating the scores made by 100 students who took a one-hour examination and a three-hour comprehensive final examination in a course in physiological hygiene. The correlation between the test and final was $r = .80$. Assuming for purposes of exposition that the final grade in the course was determined only by the comprehensive examination, the fifteen percent of the students who received less than grade C failed to do satisfactory work. It is of interest to calculate the effectiveness of the short test in selecting the satisfactory students. Suppose the poorest 15 percent in terms of the short test scores are discarded (failed). This procedure will eliminate all individual points to the left of the vertical 85 percent line. Now there were 85 successes in the unselected group. By simply counting the points in the figure it is apparent that five of these (i.e., the points to the left of the 85 percent line but above the horizontal line P-P') will be discarded; these are test "misses." The proportion of success in the retained (selected) group will therefore be 85 minus 80, divided by the total 85 who were retained. The result is 0.94, or 94 percent success in the selected group. The results with other degrees of selection appear in Table I. It is necessary to select the best 60 percent (ie., reject 40 percent) in order to eliminate all the individuals who will fail in the criterion situation. But rejecting only 10 percent achieves considerable purification of the selected group; this yields 92 percent success compared with only 85 percent without any selection.

TABLE I

PERCENT SUCCESS IN SELECTED INDIVIDUALS IN A CRITERION SITUATION INVOLVING A HIGH PROPORTION OF SUCCESS IN THE UNSELECTED GROUP (85 PERCENT) AND A LOW PROPORTION (15 PERCENT).*

Proportion Selected	*Calculation*	*Success in Selected Group*	*Calculation*	*Success in Selected Group*
100%	.....	85.0%	.....	15.0%
90	(85-2)/90	92.2	(15-0)/90	16.7
80	(85-8)/80	96.2	(15-0)/80	18.8
70	(85-17)/70	97.1	(15-0)/70	21.4
60	(85-25)/60	100.0	(15-0)/60	25.0
50	(85-35)/50	100.0	(15-0)/50	30.0
40	(85-45)/40	100.0	(15-0)/40	37.5
30	(85-55)/30	100.0	(15-1)/30	46.7
20	(85-65)/20	100.0	(15-3)/20	60.0
10	(85-75)/10	100.0	(15-7)/10	80.0

*The validity of the test is $r = .80$. Data from 100 individuals plotted in Figure 1.

In the example just given, there was a relatively high degree of criterion success (85 percent) in the unselected group. Consider now the situation if criterion success is low. The same data may be used to examine this situation if it is considered that only students who made grade A in the final examination are truly successful. Only students above the line Q-Q′ in Figure 1 are successful by this standard. If 15 percent of the unselected students are discarded and 85 percent selected by the short test, the proportion will be 15 minus zero divided by 85, or 0.176, i.e., 17.6 percent success as compared with 15 percent in the unselected group. In this circumstance, the test is without value. As shown in the right-hand portion of Table I, high success in the selected group will be achieved only if a very high proportion of the students are discarded. The results may be summarized as follows: The short test having a validity of $r = .80$ is quite effective for selecting students who are "C" or better, but is practically worthless for selecting "A" students. Furthermore, it is wasteful to discard too many students for "C" selection, but it is essential to discard a large proportion if good results are to be secured when testing for "A" students.

Another possible use of the short test might be comparable to the counseling problem mentioned earlier, using the test as a basis for advising students to drop the course. In the first case, grade C or better represents success as before. If fifteen percent are discarded on the basis of the test, there will be 10 criterion failures (i.e., those who are below the line P-P′ and to the left of the vertical 85 percent line). Ten failures in fifteen discards is 67 percent, which does not seem to be a very high proportion of failures when it is realized that 33 percent of these discarded individuals would actually be successful, if allowed to continue the course. In the case of discarding 40 percent, which served to purify completely the "accepted" group, the results are very bad indeed from

TABLE II

PERCENT FAILURES IN REJECTED INDIVIDUALS IN THE CRITERION SITUATIONS OF TABLE I.

Proportion Selected	*Calculation*	*Failure in Rejected Group*	*Calculation*	*Failure Rejected Group*
100%	.....	.....	.....	.....
90	8/10	80.0%	10/10	100.0%
80	12/20	60.0	20/20	100.0
70	13/30	43.3	30/30	100.0
60	15/40	37.5	40/40	100.0
50	15/50	30.0	50/50	100.0
40	15/60	25.0	60/60	100.0
30	15/70	21.4	69/70	98.6
20	15/80	18.8	77/80	96.2
10	15/90	16.7	83/90	92.2
0	15/100	15.0	85/100	85.0

the counseling point of view, for here there will be only 38 percent failures in the rejected individuals, and 62 percent would in fact be successful in the course if allowed to continue.

This result has occurred because, in the situation that involves a reasonably high proportion of unselected success in the criterion, the "test misses" fall into the rejected group, whereas if the proportion of unselected criterion success is relatively low, the "misses" fall into the retained or selected group. This accounts for the poor results in attempting to select the "A" students by the short test. The conclusion may be confirmed by inspection of Table II, which shows that students rejected by the test can be advised with some confidence that they will fail to secure an "A" grade. Summarizing the implications, it may be stated that when a test is used for counseling purposes, individuals advised to avoid a profession or trade on the basis of poor aptitude test scores will be poorly advised if the criterion situation involves a relatively high degree of success, because in this circumstance there will be a large proportion of "test misses" in this rejected group, even with tests of high validity. However, if the criterion situation involves a low proportion of success among unselected individuals, "test misses" will be rare and such counseling can be justified.

SELECTION-EFFICIENCY CURVES

Five members of the physical education faculty were asked to rate major students as to their professional competence. A five-point rating scale was used. Forty-eight men were rated by all five of the judges. These data will serve to illustrate the application of selection-efficiency curves.

The ratings of a single one of the judges will be used as a selection test. There will be three criteria:

(*A*) Agreement by the other four judges that a particular individual is highly competent.

(*B*) Averaged ratings of these four judges indicating that an individual is above or below an arbitrary standard of medium competence.

(*C*) Agreement of the four judges that an individual is above a minimum acceptable level of professional competence.

Biserial correlations relating the ratings of the first judge (the test scores) to success or failure in passing the three criteria have been calculated; the coefficient is $r = .76$ for (*A*), $r = .71$ for (*B*), and $r = .75$ for (*C*). The degree of success of the unselected group in criterion (*A*) was 18.8 percent, in criterion (*B*), 56.3 percent, and in criterion (*C*), 83.3 percent. These data have been used to enter each of the eleven Taylor and Russell selection-ratio tables in order to secure the information needed to construct the theoretical curves shown in Figure 2. In order to check the applicability of these tables to the present situation, actual experimentally determined values for the different

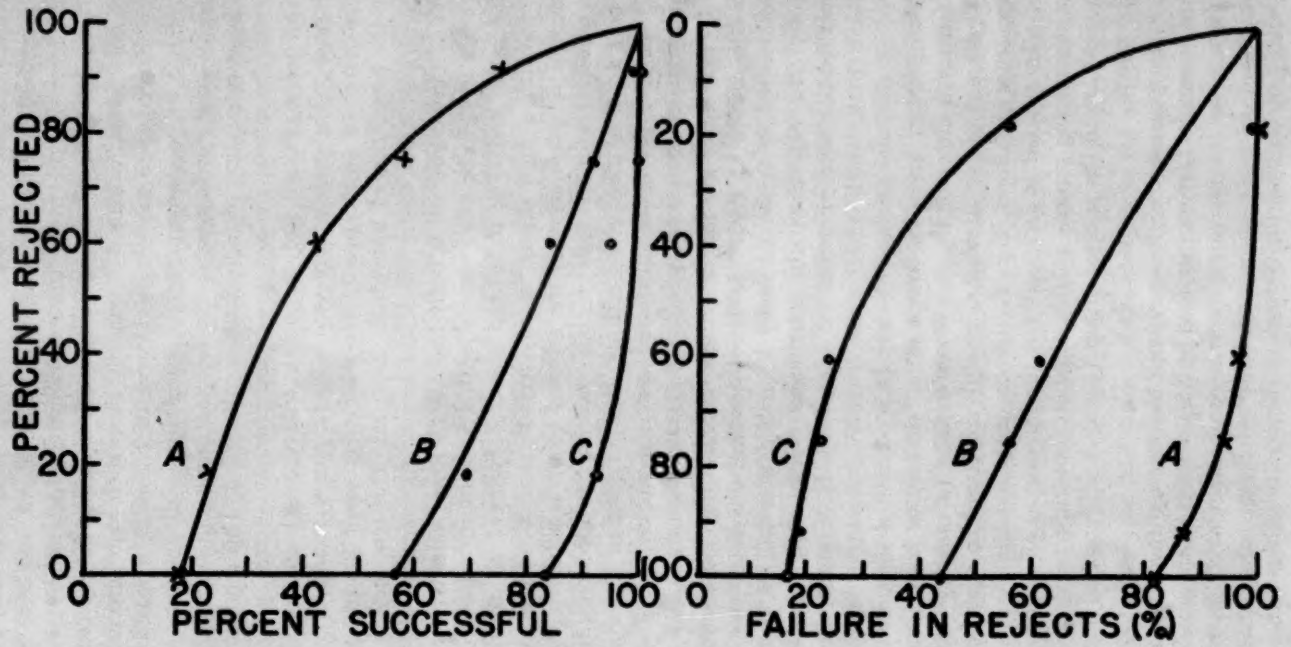

FIG. 2 (left). Amount of criterion success in a selected group with a difficult criterion (*A*), a moderate criterion (*B*), and an easy criterion (C) in relation to the proportion of individuals discarded by the test. The smooth curves are theoretical; the plotted points were experimentally determined. FIG. 3 (right). Amount of criterion *failure* in the rejected group from Figure 1. Note that the vertical axis has been inverted; purity of the rejected group can be attained only if a small number are rejected. The validity coefficients in both figures ranged from $r = .71$ to $r = .76$.

degrees of selection afforded by the rating scale have been calculated by the method described in connection with Table I, and have been plotted in Figure 2 (the fit of the experimentally determined points to the theoretical curves is exceptionally good; it cannot be expected that data on only 48 individuals will typically show such high agreement).

Referring to the left-hand curve of Figure 2, it may be seen that it tends to be comparatively vertical for small amounts of selection. Really good purification of the selected group can only be expected if a large proportion of individuals are discarded. This tendency is characteristic of curves for low unselected criterion success and would have been more striking had the unselected success been less than in this example. The middle curve tends to be almost straight, implying that the degree of purification of the selected group is linearly proportional to the proportion of individuals discarded. This tendency is characteristic of a situation where unselected success approximates 50 percent and the validity coefficient is fairly high. With lower validity, the 50 percent curves tend to be more vertical with a knee of right-hand curvature at the upper end. The right-hand curve of Figure 2 clearly demonstrates that with a high degree of unselected criterion success, the greatest amount of purification of the selected group is achieved by discarding relatively few individuals; in other words, the selection efficiency here obeys a law of diminishing returns. This is typical of situations involving 75 percent or higher unselected criterion success, particularly if the validity coefficient is fairly high (i.e., greater than $r = .60$).

For use of the judges' ratings in counseling, the essential information is shown in Figure 3. The theoretical curves in this figure were constructed from the tables by a method that will be described presently; the experimentally determined points were calculated by the method described in connection with Table II. Consider first the right-hand curve, which concerns the criterion involving a high standard of success that was attained by only 18.8 percent of the unselected group. Men who are given the lowest rating are sure to fail this criterion and men given either of the two lowest ratings (a category which includes 60 percent of the unselected group) are found to show a 97 percent incidence of failures. The effectiveness of the test improves as the number of men rejected is decreased, but the increase in efficiency obeys a law of diminishing returns. In this situation counseling is effective; the "test misses" are in the *selected* group (curve A of Figure 2) rather than in the rejects.

In contrast, the efficiency with the criterion involving a low standard of success that was attained by 83.3 percent of the unselected group obeys a law of *increasing* returns as the number of men rejected is decreased; however, the effectiveness never gets high enough to justify the use of the test for counseling the rejects in this situation (except perhaps in the case of the poorest three or four percent of the unselected

men). This result occurs even though the rating test had a validity of r = .75, considerably higher than is usually encountered in aptitude tests. The reason for the poor results is of course explained by the fact that the "test misses" are now in the rejected group; this is why the test was effective in purifying the *selected* group with respect to success in the 83.3 percent criterion (curve C, Figure 2).

When the rating test is used in connection with criterion situation B (56.3 percent of unselected success), the efficiency is approximately equal with respect to counseling the rejects and improving the selected group; here, the test misses are divided almost-equally between the selected group and the rejects. Furthermore, efficiency tends to be a linear function of the smallness of the selected or rejected group (middle curves of Figures 2 and 3).

Comparison of Figures 2 and 3 shows that the most effective placement of a "cutting" score is very different, depending on whether the test is to be used for improving the selected group or counseling the rejected group. In the case of the 18.8 percent or low-success criterion (curves A), improvement of the selected group requires a high degree of selection; efficiency *increases* rapidly in the upper regions of the curve. But if the test is used for counseling the rejects, a relatively low cutting score is best, because the efficiency increases very slowly in the upper regions of the curve.

In the case of the 83.3 percent or high success criterion (curves C), a fairly low placement of the cutting score is desirable if the test is being used to improve the selected group, because efficiency changes very slowly in the high regions of the curve. On the other hand, if the test is being used for counseling the rejects, it is absolutely worthless unless the cutting score is placed very high; efficiency increases rapidly in the upper regions of the curve. It seems relevant to mention at this point that, in general, counseling of rejects is concerned with criterion situations in which unselected success is well above 50 percent, and curve C is probably a fairly typical curve.

Several other examples of the agreement between the theoretical curves and experimental results are shown in Figure 4. In connection with a study of exercise at high altitude,[2] 26 men were given a synthetic 38,000 ft. "flight" in a decompression chamber while engaging in muscular activity. Subsequently the same men were given another "flight" with a different exercise. Self-ratings of the severity of the altitude pain experienced by these men on their first exposure may be used as altitude tolerance scores; individuals who had no symptoms of decompression sickness during the second exposure may be considered successful in their second or validating flight, while the men who had symptoms may be considered as having failed the criterion. The unselected men showed 23.5 percent criterion success. The biserial correlation between test

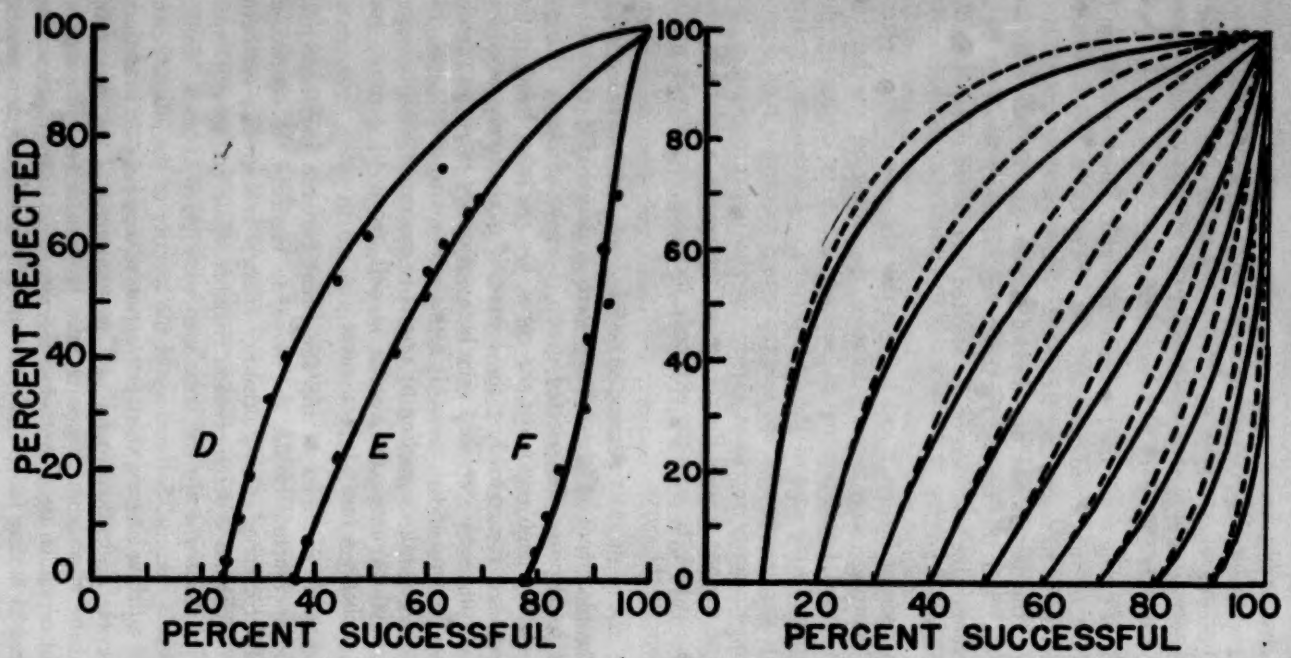

FIG. 4 (left). Improved performance at high altitude resulting from selection by a test having a validity of r = .64 (curve *D*), r = .68 (curve *E*) and r = .52 (curve *F*). The smooth curves are drawn from theory; the plotted points were experimentally determined. FIG. 5 (right). Theoretical selection-efficiency curves for r = .70 (dotted lines) and r = .80 (solid lines). These curves predict success in the selected group; by inverting the vertical axis so that 100 is at the bottom and zero at the top, and relabeling the horizontal axis to read *percent failure*, they will predict failure in the rejected group.

scores and criterion results is r = .64. This information, with the aid of the tables, makes it possible to construct curve *D*.

In another high-altitude experiment, 148 men were "flown" at 33,750 ft. and scored as to the time of the first appearance of altitude pain. The presence or absence of symptoms of decompression sickness in a subsequent exposure was noted. The unselected criterion success was 36.6 percent, and the validity coeffcient was r = .68. These data yield the middle curve of Figure 4. The third altitude experiment, on 134 men, made use of pain intensity scores in the first "flight." Failure in the second or criterion flight was defined as the occurrence of intolerably severe symptoms. The unselected group showed 77.2 percent success by this criterion; the validity coefficient was r = .52. Curve *E* shows the efficiency of the test. In all three of the experiments, the actual occurrence of criterion success at various degrees of selection has been calculated as in the Table I data, and plotted on the graph.

It will be noted that the agreement between theory and experiment is exceptionally close in the case of curve *E* of Figure 4. This is a very interesting observation because the test scores in this example were highly skewed, in contrast to the relatively normal distributions of the scores in *D* and *F*. Comparison of the agreement between theory and experiment in over a hundred curves similar to these[1,4] indicates that the method can be used with skewed distributions, at least for validity coefficients smaller than r = .80, provided that biserial correlations are used. Such usage is of course a violation of the assumption of a normal bivariate correlation surface although the statistical theory is based on this assumption.[6, 7] Practically, the most important assumption appears to be the postulation that the correlation between test and criterion is uniform throughout the range of scores. On the whole, the theory appears to underestimate slightly the amount of improvement in the selected group.

THEORETICAL TABLES AND CURVES

Most accurate are the Pearson tables[6] of the volume of the normal bivariate surface, familiar to statisticians for their application in calculating positive and negative tetrachoric correlations. In the present application, Pearson's Table VIII may be used for estimating improvement in the selected group and Table IX may be applied to the rejects. Unfortunately, these tables are not arranged in the most useful form for the present purpose; furthermore they will not be very intelligible to most non-mathematicians. The Taylor and Russell selection-ratio tables[7] are easier to understand and have adequate accuracy for most purposes.

The writer has prepared a set of tables[3] of the form shown in Table III that will be found useful if the problem of testing is visualized along the lines of the present paper. An advantage of this form is that a single table serves to give the purity of both the *selected* group and the *rejects*. To illustrate: consider the situation for r = .60 with 20 percent

unselected failure in the criterion. Enter the 20% column (third from the right) to find the percent failures in the selected group as a function of the percent of unselected men discarded by the test (extreme left-hand column). For example, if the 10 percent of individuals making the poorest test scores are discarded, there will be 16 percent failures, in the selected group; if 20 percent are discarded, there will be 13 percent failures; if 30 percent are discarded, there will be 10 percent failures, and so on.

In the unselected group, with 20 percent success there were 80 percent failures. Enter the 80% column, but this time from the top downward and read 40 percent *success in the rejects* if 10 percent of the unselected group is discarded, 50 percent success in 20 percent discards, 57 prcent success in 30 percent discards, and so on. For any entry in the body or margin of this table, a result given in percent success can be converted to percent failure by subtracting from 100, and a result given as percent failure can be converted to percent success by subtracting it from 100.

TABLE III

(From Henry, 1946[3])

TEST EFFICIENCY FOR r = .60

PERCENT FAILURES IN THE SELECTED GROUP AS A FUNCTION OF THE PERCENT OF THE UNSELECTED GROUP THAT IS DISCARDED.*

90%	(10%)	61	40	26	17	10	6	3	1	0	0
80%	(20%)	70	50	36	25	16	10	5	2	1	0
70%	(30%)	75	57	42	31	21	13	8	4	1	0
60%	(40%)	79	62	48	36	25	17	10	5	1	0
50%	(50%)	82	66	53	40	30	20	13	6	2	1
40%	(60%)	84	70	57	45	34	24	15	8	3	1
30%	(70%)	86	73	60	49	38	27	18	10	4	1
20%	(80%)	88	76	64	52	41	31	21	13	5	2
10%	(90%)	89	78	67	56	46	35	25	16	7	3
0%	(100%)	90	80	70	60	50	40	30	20	10	5

*The table also gives the percent success in the discarded group as a function of the percent of the unselected group that is discarded (use figures enclosed in parentheses).

It has also been found useful to prepare a set of curves for each magnitude of correlation coefficient by 10's. Such a family of curves is shown in Figure 5. Graphic interpolation can readily be made for any amount of unselected success, while the use of two magnitudes of correlation in each graph makes it easy to interpolate between a range of ten points in correlation. Confusion may be reduced by lightly shading with a pencil the areas bounded by each pair of lines. These graphs are not very accurate for extremely high levels of rejection, but the tables can be resorted to for the infrequent occasions when this region is being used. It may be noted that by simply relabeling the horizontal

axis of Figure 5 to read "Percent failure in rejects" and inverting the numbers on the vertical axis so that 100 is at the bottom and zero percent is at the top, the curves may be used to predict results with the rejected group.

Inspection of Figure 5 with respect to the horizontal distances bounded by the pairs of lines reveals that alteration of a validity coefficient by as much as ten points does not result in a very great change in test efficiency in a considerable portion of the possible situations. Variations in validity have a modest effect when unselected success is high; for medium amounts of unselected success, altering the validity coefficient has a considerable effect if half or more of the unselected individuals are discarded by the test. If unselected success is low, the effect is large when a large proportion is discarded. These observations apply to purification of the selected group. In the case of rejects, alteration of the validity coefficient has the greatest effect when unselected failure is high; when less than half of the unselected individuals are rejected if unselected failure is in the region of 50 percent; at *low* levels of rejection when unselected failure is infrequent.

When a test is used for practical purposes, its true validity in the particular application is seldom known; similarly, the true proportion of unselected criterion success will not often be known with much accuracy. Nevertheless, some reasonable approximation of these statistics must be made if use of the test is to be justified by any method of evaluation. Using such approximations it should always be possible to decide at least whether to use the right-hand, middle, or left-hand parts of Figure 5 (or a similar figure for some other range of validity coefficients). This will provide a rational basis for correctly placing the "cutting" score, and, equally important, will serve to prevent injudicious placement. If the cutting score is fixed by policy, as would be the case, for example, if a fixed number of professional students were accepted from all applicants for training, or if it were customary to fail a certain proportion of students in a large class, it is a simple matter to estimate the *reduction* in failures fairly accurately even if unselected success and validity are approximations. This is possible because adjacent curves in diagrams such as Figure 5 tend to be similar in form. Obviously the above arguments apply equally well if interest is centered in the counseling situation.

SUMMARY AND CONCLUSIONS

Analysis of a scatter diagram of individual test scores plotted against criterion results reveals that the efficiency of a test varies widely depending on the proportion of individuals selected by the test and the unselected incidence of success and failure, even though the validity coefficient remains constant. If the amount of unselected success is high and a comparatively small proportion of individuals are selected, "test

misses" tend to fall in the rejected group, making the test efficient for improving the selected group and inefficient in purifying the rejected group. However, if the proportion of unselected success is low, the "test misses" tend to fall in the selected group. This results in low test efficiency insofar as the selected group is concerned, but efficiency is high with respect to purifications of the rejected group. When unselected success and failure are about equal, the "test misses" tend to be distributed equally between the selected and rejected groups.

Using the theoretical volumes of the normal bivariate correlation surface, *selection-efficiency* curves may be constructed for specific proportions of unselected success and magnitudes of test validity. These curves are found to agree very closely with experimentally determined test results. They may be arranged to give predictions of criterion success (or failure) in the selected group resulting from various percentages of selection, or to give predictions of success or failure in the rejected group. For practical purposes, it is found convenient to use superposed curve-families for two magnitudes of correlation coefficient. This technique permits rapid interpolation of both correlation and criterion success with sufficient accuracy for most applications. Simplified test-efficiency tables are also valuable; an example of such a table is shown. The method is useful for test evaluation even though validity and unselected success are only approximately known.

Study of the selection-efficiency curve system leads to the following generalizations:

1. The incidence of criterion success and failure in the selected and rejected groups can be determined for any specified cutting score on a test, if the validity coefficient and unselected incidence of success or failure are known.

2. For any particular magnitude of the validity coefficient, a large improvement in the selected group results from discarding relatively few individuals, and further discarding obeys a law of diminishing returns, when the incidence of unselected criterion success is high.

3. When there is a low incidence of unselected success in the criterion, improvement obeys a law of increasing returns, but a large proportion of the unselected group must be discarded before there is much improvement in the selected group.

4. When unselected success is low, there is a high degree of purity of the rejected group when relatively few individuals are discarded by a test, but the rejects become mixed with an increasing proportion of "misses" as more and more individuals are discarded.

5. With high unselected criterion success and relatively few individuals discarded by the test (the most desirable situation for improving the selected group), the rejected group is extremely impure if more than a few percent are discarded.

6. When unselected success is close to 50 percent, improvement of the selected group tends to be linearly proportional to the percent of individuals discarded. In contrast, the rejected group tends to become increasingly impure in linear proportion to the percent discarded.

7. Alteration in the magnitude of a validity coefficient has relatively little effect on the improvement of the selected group when unselected success is low and few individuals are discarded; the effect is greatest when over half are discarded and unselected success greater than 50 percent. On the other hand, the purity of the rejects is most affected by alteration in validity when unselected success is low and relatively few individuals are discarded.

8. All these effects are small at low levels of validity, but become quite important for validity coefficients that are large enough to justify the use of a test in a practical situation.

BIBLIOGRAPHY

1. Henry, F. M., "Altitude Pain," *J. Aviat. Med.* 17:28-55 (1946).

2. ....., "The Role of Exercise in Altitude Pain," *Am. J. Physiol.*, 145:279-284 (1946).

3. ..... *Statistical Methods and Tables of Test Efficiency.* Berkeley: University of California, 1946 (Mimeographed—available on request).

4. ..... and A. C. Ivy. "Preselection Tests," in *Decompression Sickness*, J. F. Fulton, ed. In press, 1947.

5. Johnson, H. M., "Maximal Selectivity, Correctivity and Correlation Obtainable in 2x2 Contingency-tables," *Amer. J. Psychol.* 58:65-68 (1945).

6. Pearson, K. *Tables for Statisticians and Biometricians*, Part 2. London: Cambridge University Press, 1931.

7. Taylor, H. C., and J. R. Russell, "The Relationship of Validity Coefficients to Practical Effectiveness of Tests in Selection," *J. Appl. Psychol.*, 23:565-578 (1939).

Hidden Possibilities For Research In Physical Education and Recreation

By Florence Stumpf and Frederick W. Cozens
University of California
Berkeley

(*Submitted for publication, January, 1947*)

THE original impetus for writing this article might be traced to the reading of a book, a book that formed just another link in that seemingly endless chain of books one should read to "keep up" with the best educational thought of the day, a book that came into being as the result of the work of a committee appointed by the American Council of Learned Societies.[1*]

This committee sounded a call to "specialists," working in their own restricted fields, to attempt, from the point of view of their own interest, to achieve an "historico-philosophical" orientation; to awaken to the importance of setting their specific problem in a larger context, and thus to avoid meaningless particularity. While admitting that it is natural for man to orient himself first to his own culture, and his immediate geographic environment, they maintained that a liberal education should lead him to discover that his own culture has been powerfully influenced by other cultures in other portions of the globe. It should bring to him an awareness of the existence of other cultures and a desire to explore their distinguishing characteristics and their historical development. Finally such education should bring him also an awareness of the many factors, economic, social, political, geographic, and religious, which have been involved in the differentiation of one culture from another. The authors deplored the almost exclusive preoccupation, in formal education, with our own culture to the neglect of the rich and ancient cultures of the Near East and Far East. They pointed to the global thinking generated by the recent war and welcomed such thinking as a powerful corrective to this limitation.

This set us to thinking, most seriously, of our own "specialty" and led us to attempt an objective examination of the literature in our field, to see how it would stand up against the challenge of provincialism, in view of this concept of a world-wide orientation.

The results, in terms of information available on what has happened in relation to the use of our particular educational tool outside

---

This manuscript is intended as a preface to a series of articles illustrating the possibilities as outlined in the paper. Though it is not research in the strict sense, it is published as an introduction to research. The first article will deal with the role of games, sports, and recreational activities in the cultural pattern of a primive people.

*Superior figures refer to numbered bibliography at end of article.

that part of the globe which might be called the domain of Western civilization, were disappointing indeed. Scattered bits of information on the work of the Y.M.C.A.'s and Y.W.C.A.'s, fragmentary comments of occasional visitors, a few fine pieces of work done by citizens of other parts of the globe, studying in the schools of our country, make up the sum total of the results.

And yet, stepping out of the confines of our own literature, our search has led us to an untold wealth of material collected by earnest scholars and investigators in other fields, fascinating, challenging, provocative material, awaiting the attention of students with initiative, imagination, and tenacity. The course is not charted, true, but enough guides may be found to keep the ambitious adventurer from losing his way.

The archives of research material in ethnology and anthropology alone contain enough to engage the energies of interested students for years to come. The most cursory examination of the number of fine studies done by ethnologists and anthropologists on the play life, games, amusements, and ceremonial dances of primitive peoples, will convince the most skeptical, unless he be utterly lacking in perception, of the possibilities inherent in such studies with relation to our own field.

A study of the literature dealing with the recreational life and physical activities of the peoples of the Pacific will provide for the physical educator and the recreation worker unexcelled laboratory material in social phenomena.

Is it not possible in our search for the answers to the "worthy use of leisure time" that we can learn something from the Polynesians? Keesing refers to them as having "the highest standard of leisure the world has ever known."[2] The recreational activities with which these people occupied their leisure, and the contributions these activities made to their total living, conform to many standards which we set up for our own programs. It was recreation for all, not for a few. Activities were closely integrated with actual living. Participation was the role of each individual in his own small way, and the effect was a powerful socializing influence tending to promote group solidarity and establish a definite feeling of "belonging" in all.

The importance of the social value of inter-village and inter-community competitive games was very great. Young and old were brought together as a group and all efforts were made to make each competition a truly festive occasion. It was an obligation on the part of the local people to provide food for the visitors, spectators as well as competitors, and when the games were played at another village this hospitality was repaid and amplified if possible. Polynesians attending the sports events of people of a so-called higher culture are struck by the lack of organized hospitality in providing free food and drink, a custom so characteristic of their own culture.[3]

Surely it is no less the business of the teacher of physical education and the recreation worker than it is that of the anthropologist to study human society and culture, to learn the way in which human beings interact in society, to learn the nature and function of social organizations, and means and effects of culture contact and change. He will learn that specific physical activities and recreational pursuits vary from society to society, but that such activities are common to *all* societies and serve certain constant functions. Recreational pursuits held in common by a social group serve to enhance the social consciousness of that group and also to give psychological security to the individual. Malinowski[4] has said that such forms of relaxation and artistic stimulation of the nerves and muscular system are a condition of healthy communal life and also fertilizing factors in cultural development and progress.

Before proceeding we feel it is only fair to post certain storm warnings. In our judgment certain prerequisite qualities should be present in the hopeful research worker in this field. First he must have sufficient sociological understanding to realize that "man does not live by bread alone"; and that after the basic needs of food, clothing, and shelter, and that other need of humans to explain the imponderable mysteries of life and death have been met, there still remains the surplus of time and energy to be spent on the things which give life joy and color and meaning.

He should possess the fundamental conviction that play is as universal a part of the cultural patterns of the world, primitive or modern, as the economic, familial, religious, or political life. He should realize further, that to minimize, ignore, or pervert this fundamental human need is dangerous. We need not search far in our own culture to find evidence that when children are allowed to proceed along the road to maturity without proper training and guidance for the pursuit of their own recreational life, both while they are children and after they become adults, we have a multitude of social ills that are much more difficult to correct than they might have been to prevent. Among adolescents we have increasing juvenile delinquency and an unhealthy preoccupation with commercialized forms of recreation which do not provide that outlet for creative energy that comes from active, interested participation. Among adults we have lethargy, boredom, neuroses, and over-indulgence in many undesirable forms.

The student who would attempt this sort of research must be free from what a well-known anthropologist [5] has termed "ethnocentrism," a form of cultural myopia that limits the mental vision and cultural horizon to the confines of one's own society. He must have a sufficiently developed understanding to comprehend and appreciate that man in his varied career in this world has worked out an endless variety of adaptations to life's problems, and that through a sympathetic

comprehension of the answers found to these problems in other societies we may be in a better position to understand our own society, its contradictions and failures, its achievements and its values.

Here is a challenge to students and teachers who are vitally interested in games, sports, recreational activities, as an integral part of the culture of all peoples; who are eager to seek out and understand the basic sociological significance of the aspect of culture:[6] students who are convinced that all human activities are interrelated, and that none is so trival that it may not influence or illustrate others which may seem more important; that while the food they eat and the tools they use may determine the chances for survival of a people, the games they play or the stories they read may establish the character of a nation.[7]

Here is research to be done by individuals who realize that understanding between peoples forms the basis of our hope for the future of civilization,[8] and who have the firm conviction that the tools of our profession in the hands of skilled and sincere teachers may contribute in full measure to the building of such an understanding.

Such research needs students who are curious, who want to seek out the significance of such a fact as is represented in the knowledge that the National Council of Physical Education in Chile is placed under the immediate control of the Ministry of War, while in Bolivia the General Bureau of Physical Education is under the control of the Minister of Public Instruction. Is the difference in control one of administrative expediency or is it evidence of a different philosophical approach?

What part has instruction in physical education activities played in the "cultural mission" experiment in education in Mexico? Do we find in this experiment any implications for the solution of our problems in rural education, particularly in economically depressed areas? Could this be the answer to instruction in physical education in our many thousands of one-room, one-teacher schools?[10]

What part is our profession and its thousands of members to play in the future international role of the United States Office of Education?[11] How are the tools of our profession viewed by the United Nations Organization?[12] How will physical education activities and preparation for "worthy use of leisure" enter into the plans of UNESCO? Because we believe that physical education should form an integral part of any scheme of education, and, in its broader aspects including recreation and rehabilitation, an important instrument in the total picture of civilization, the members of our profession will follow with great interest the deliberations of the representatives at UNESCO.

The possibilities of certain types of historical studies in our field have been briefly suggested. If a complete history of physical educa-

tion is ever to be written, we must take cognizance of the vast geographical areas of the world which as yet are relatively untouched in our literature. The historical record of the recreational life and physical activities of the many millions of human beings outside of our western civilization cannot be ignored if we hope to be able to trace the intricate interrelationships of our field with other social phenomena.

BIBLIOGRAPHY

1. *Liberal Education Re-examined, Its Role in a Democracy.* Committee appointed by the American Council of Learned Societies: Theodore M. Greene, chairman, Charles C. Fries, Henry M. Wriston, William Dighton. New York: Harper & Bros., 1943.

2. Keesing, Felix M. *The South Seas in the Modern World.* New York: The John Day Co., 1941.

3. Buck, Peter H. (Te Rangi Hiroa) *Samoan Material Culture.* Bishop Museum Bulletin, No. 75. Honolulu: 1930.

4. Malinowski, Bronislaw. *The Dynamics of Culture Change.* New Haven: Yale University Press, 1946.

5. Beaglehole, Ernest, "New Zealand Anthropology Today," *Journal of the Polynesian Society,* 46 (September, 1937), pp. 154-172.

6. Roucek, Joseph S., and associates. *Sociological Foundations of Education.* New York: Thomas Y. Crowell Co., 1942.

7. *The Interpretation of History.* Joseph R. Strayer, ed. Princeton, N. J.: Princeton University Press, 1943.

8. Keesing, Felix M. *Native Peoples of the Pacific World.* New York: The Macmillan Co., 1945.

9. Smith, Henry L., and Harold Littell. *Education in Latin America.* New York: The American Book Company, 1934.

10. "Education for Social Control," *Annals of the American Academy of Political and Social Science.* November, 1935.

11. "International Frontiers in Education," *Annals of the American Academy of Political and Social Science,* 235 (September, 1944),

12. *Education and the United Nations.* Report of a Joint Commission of the Council for Education in World citizenship and the London International Assembly. Washington, D. C.: American Council on Public Affairs, 1943.

A Factor Analysis of Some Cardiovascular-Respiratory Variables and Tests

By LEONARD A. LARSON
New York University
New York City

(Submitted for publication, January, 1947)

THREE studies have been reported on the analysis of the interrelationships of cardiovascular-respiratory variables. Two of these studies were on college men[3, 8*] and one on college women.[7] The purpose was to analyze the table of intercorrelations using the Thurstone Factor Analysis Method[9, 10 11] to determine the principal components constituting the table of intercorrelations or the components describing the cardiovascular-respiratory variables. When dealing with a number of variables one can assume that they are all unrelated and therefore there are as many abilities as tests; or one can assume that they are all perfectly related and therefore there is only one ability or characteristic. Neither of these hypotheses is correct as indicated by the reported studies and some number between these extremes describes cardiovascular-respiratory function.

McCloy[3] found in analyzing 14 cardiovascular variables that they were described by at least eight principal components. The variables included horizontal and standing systolic, diastolic, and pulse pressures; a comparison of pulse rate and pressures between the horizontal and standing positions; and pulse rate after exercise. Three of the isolated factors were identified as factors of pulse rate (heart rate, standing heart rate, and heart rate return to normal), and five as factors of blood pressure (general blood pressure, pulse pressure, reclining diastolic pressure, effect of position on systolic, and reclining systolic pressure).

Utilizing 32 cardiovascular variables in examining a group of college women, Murphy[7] found nine primary components. Three of the isolated factors were described as factors of pulse rate and six as blood pressure factors.

In a study conducted by Quamme,[8] 20 cardiovascular-respiratory variables were analyzed by the factor method. Eight factors were isolated. Four were described as factors of pulse rate (circulatory recovery to exercise, effect of pulse rate to exercise, pulse rate return after exercise, and pulse rate response to postural changes), and four as factors of blood pressure (head or maximum pressure, diastolic changes with body

The data for this study were gathered in the course of research conducted under the auspices of the National Research Council Committee on Selection and Training of Aircraft Pilots with funds made available by the Civil Aeronautics Administration. The author wishes to acknowledge help given on statistical computation by Mr. H. J. Quamme.

*Superior figures refer to numbered bibliography at end of article.

position, diastolic pressure, and maximum pressure changes with body position).

STATEMENT OF PROBLEM

The purposes of this research are two-fold (1) to analyze some cardiovascular-respiratory variables for their principal components, and (2) to determine the physiological characteristics of ten selected cardiovascular test index scores on the basis of the isolated factors.

PROCEDURE

Experimental.—Nineteen cardiovascular-respiratory variables were administered to 145 Springfield College men, the majority of whom were major students in health and physical education. Some of the variables were used to score ten cardiovascular tests. The total of 29 variables and tests was used in the statistical analysis. The nineteen variables served as material for the isolation and identification of the factors. The ten test batteries were included to determine their physiologic characteristics on the basis of the isolated factors. The following precautions were used in obtaining data for the statistical analysis: (1) exercise was not allowed before the examination (during that day); (2) subjects with emotional disturbances were not used; (3) subjects not receiving a normal amount of sleep the night previous to the examination were eliminated; (4) subjects with heart defects (medical examination) were not used; (5) subjects having a heavy meal just before the examination were asked to return in one hour; (6) subjects with colds were not used; and (7) the subjects selected represented a physically fit group of young men (18-24 years of age), all of whom participated in physical education activities as part of their education. The measurements were all made in one week in September. All questionable results were rechecked and were dropped or retained on the basis of the above criteria.

Statistical.—The Thurstone Method of Factor Analysis[10,11] was used in the isolation of the factors. This method begins with a table of intercorrelations (406 rectilinear correlations in this problem) and results in the isolation of a number of *factors* or *components* which constitute the major part of the tests analyzed and gives a correlation between the factor and the tests used in the analysis. The method assumes that the tests and variables are related and that they can be described by a relatively small number of factors. The name given to the factor, or identification, is determined by how much one knows about the tests and the field within which he works. The factor method involves no assumption as to the nature of the factors. In this research two physiologists have reviewed the statistical results and have given suggestions as to the nature of the factors isolated.

Limitations.—In the interpretation and application of the results of this research, the following should be considered:

1. The variables and tests selected cannot be assumed to indicate a complete measure of cardiovascular-respiratory functions.

2. The isolated factors and their identification are based on the variables (not test batteries) used in the study and on the judgment of physiologists.

3. The sample consists of Springfield College men from 18 to 24 years of age, all of whom participated in physical education activities.

4. The factor analysis studies reviewed cannot be directly compared because the variables are not identical, the system of rotation and the identification of the factors involve judgment skill, that is, the best rotation may not be made and the factors may be identical but given different names, and the data are on different institutional populations and on men and women.

5. The variables and tests are not all completely described by the isolated factors. Factors will appear only if they are present in at least two of the tests analyzed. Because some of the tests are not completely described by the isolated factors, the implication is that other factors are possible.

6. The ten test batteries are described by the isolated factors in terms of their "index scores."

7. The reliability range of the cardiovascular-respiratory variables and tests is .70 to .95 which means that a high degree of confidence as to the stability of the factor cannot be made in some instances.[4]

RESULTS

The results of the factor analysis of the table of intercorrelations (Table I) show that there is no common factor, that is, a factor common to all variables and tests used in this analysis (Table II). The eight isolated factors are described by independent clusters of variables, which give evidence that cardiovascular-respiratory variables are sufficiently independent or unrelated so that *at least* eight factors are necessary to describe cardiovascular-respiratory function. The eight factors cannot be considered entirely independent or entirely unrelated, but they are not sufficiently related for cross prediction.

With the aid of two physiologists* the eight isolated factors were analyzed and physiologically described. The single variables were statistically analyzed to determine their "net" contribution to each factor. Those yielding significant "net" contributions were selected as predictive of the factor. The multiple correlations (Table III) indicate the significance of the prediction and were used as an aid in the identification of

*The writer wishes to express appreciation to Dr. P. V. Karpovich and Dr. Percy Dawson for their description of the isolated factors.

TABLE I

Zero-order Intercorrelations of Cardiovascular-Respiratory Variables and Tests

		(1) TEST HOR. P.R.											
		1	2	3	4	5	6	7	8	9	10	11	12
Std.* P.R.	2	.7461											
Hor. S.P.	3	.1075	.0125										
Hor. D.P.	4	.1862	.0992	.1635									
Hor. P.P.	5	—.1047	—.1051	.5780	—.5810								
Sit. S.P.	6	.0825	—.0099	.7789	.1780	.4409							
Sit. D.P.	7	.0165	—.0119	.1477	.6257	—.3237	.3164						
Sit. P.P.	8	.0536	—.0210	.6125	—.3097	.6544	.6993	—.4409					
Std. S.P.	9	.1346	—.0100	.7051	.2634	.3816	.7714	.3751	.4443				
Std. D.P.	10	.0620	.0033	.1519	.6226	—.3008	.2268	.6245	—.2715	.4445			
Std. P.P.	11	.0791	—.0197	.5909	—.2613	.6506	.5972	—.1521	.6899	.6424	—.3963		
V.C.	12	—.1810	—.0787	.1281	—.0009	.0500	.1600	.0788	.0894	.1095	.1001	—.0010	
B.H. after Ex.	13	—.1727	—.1682	.0155	—.0228	.0086	.0066	—.0344	.0447	.0298	.0952	—.0692	.4084
P.R. 2 Min. after Ex.	14	.7380	.7760	.0743	.1542	—.0473	.0676	.0458	.0202	.1456	.1022	.0293	—.0545
Diff. Std. P.R.-P.R. 2 Min. after Ex.	15	.2822	.0878	.1211	.0517	.1189	.1283	.0307	.1050	.2552	.0533	.1707	—.0066
Diff. Hor. Std. P.R.	16	—.1704	.5019	—.1604	—.0783	—.0594	—.1594	—.0324	—.1415	—.2355	—.0641	—.1744	.1168
Diff. Hor. Std. S.P.	17	.1071	.0022	—.1583	.1852	—.1303	.1654	.3423	—.0924	.4765	.4077	.2137	—.0239
Diff. Hor. Std. D.P.	18	—.1421	—.1088	—.0274	—.3623	.2578	.0280	.0174	.0006	.2077	.4382	—.1757	.1406
Diff. Hor. Std. P.P.	19	.2228	.1284	—.0938	.4318	—.4793	.1228	.2488	—.0551	.1949	—.0618	.3133	—.1131
McCloy Test	20	—.4615	—.7081	.0715	.3263	—.1282	.1086	.3750	—.1768	.2353	.5974	—.2679	.0623
Barach Test	21	.5981	.7371	.2870	.3533	—.0533	.3151	.3017	.0536	.2671	.4173	.1336	.0317
Stone Test	22	.1089	.0476	.3112	—.4135	.5091	.2832	—.7030	.8069	.0940	—.4258	.4825	.0377
Basal Metab.	23	.5834	.6715	.4270	—.1023	.3759	.4152	—.1162	.4778	.4347	—.2560	.7007	—.0126
Tigerstedt Test	24	.0973	.0111	.4060	—.3787	.5843	.3939	—.3044	.6122	.3782	—.5920	.8949	—.0280
P.P. & P.R. (Std)	25	.3793	.3665	.5364	—.1787	.5291	.5287	—.1384	.6097	.5966	—.3304	.8829	—.0076
P.P. & P.R. ÷ Dias. (Std)	26	.2998	.3076	.4125	—.3274	.5356	.3737	—.3052	.5917	.3957	—.5638	.8594	—.0299
Crampton Test	27	.1558	—.1748	—.0944	.1573	—.0721	.1354	.2432	—.0418	.5619	.3357	.2430	—.0640
Schneider Test	28	—.4712	—.6338	—.0400	.0500	.0396	.1476	.1711	.0114	.4760	.2706	.0654	.0567
McCurdy-Larson Test	29	—.1392	—.1416	—.0150	.0974	—.1419	.0622	.2983	—.1696	.0719	.2109	—.1584	.4596

TABLE I (cont.)

Zero-order Intercorrelations of Cardiovascular-Respiratory Variables and Tests

13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
—.1902															
—.1477	.6437														
—.0452	.2214	—.2325													
—.0835	.1236	.1831	—.1395												
.0608	—.0574	.0431	.0227	.3391											
—.1103	.1242	.0341	—.1065	.4641	—.5613										
.1397	—.5048	—.0232	—.4328	.2657	.3601	—.1279									
—.1165	.6423	.1999	.2828	.3703	.1523	.1912	—.2850								
—.0103	.0814	.0740	—.0773	—.1760	—.0195	—.1057	—.2958	—.0816							
—.1752	.6031	.1950	.2204	.1860	—.1654	.3074	—.6709	.6155	.4067						
—.1045	.0804	.0907	—.1005	.0764	—.2709	.2914	—.3862	—.0552	.5737	.6775					
—.1688	.3810	.1766	.0431	.2234	—.1696	.3322	—.4911	.3952	.5034	.9166	.8718				
—.1677	.2963	.1144	.0674	.0529	—.2919	.3030	—.5903	.1835	.5676	.8499	.9336	.9485			
—.0516	.0527	.2750	—.4641	.8657	.2749	.4105	.3605	.1962	—.1083	.0829	.1084	.1743	.0276		
.0828	—.5231	—.1359	—.3426	.4333	.2771	.0838	.6579	—.3135	—.0711	—.3691	—.0039	—.1698	—.2116	.4940	
.6000	—.1434	—.0793	—.0111	.0773	.1309	.0034	.2545	—.0687	—.2483	—.2106	—.1548	—.1921	—.2428	.0509	.1619

*Hor., Horizontal
P.R., Pulse Rate
P.P., Pulse Pressure
Sit., Sitting
Std., Standing
D.P., Diastolic Pressure
S.P., Systolic Pressure
B.H., Breath-holding
V.C., Vital Capacity
Min., Minutes
Sec. Seconds
Ex., Exercise

N = 145

N=145

TABLE II

ROTATED FACTOR LOADINGS OF CARDIOVASCULAR-RESPIRATORY VARIABLES AND TESTS

	I	II	III	IV	V	VI	VII	VIII	Rotated h^2	Unrotated h^2
1. Hor. P.R.*	.0820	.1272	.7227	.2259	—.2236	.1689	—.0403	—.2557	.7417	.7417
2. Std. P.R.	.0021	—.1955	.8287	—.1510	.1311	.1292	—.1126	—.4465	.9937	.9932
3. Hor. S.P.	.3870	.1405	.2173	—.2154	—.0440	—.2701	—.5237	.5028	.8651	.8650
4. Hor. D.P.	—.5050	.1763	.3468	—.4439	—.0182	—.2597	—.3187	.2136	.8183	.8181
5. Hor. P.P.	.6873	.0270	—.1109	—.5734	—.1006	—.0529	—.0134	.2560	.8928	.8923
6. Sit. S.P.	.3270	.0704	.2963	—.0777	—.0546	—.2087	—.4610	.6490	.8873	.8873
7. Sit. D.P.	—.5689	.0393	.3083	.0839	.2867	—.3096	.0407	.5176	.8748	.8751
8. Sit. P.P.	.7601	—.0670	.0522	—.1277	—.2780	.0292	—.4620	.2031	.9341	.9353
9. Std. S.P.	.2208	—.0381	.3735	—.0368	—.0997	—.0733	—.0639	.8471	.9282	.9288
10. Std. D.P.	—.6766	—.1094	.2792	—.0857	—.3573	—.0917	—.1502	.4539	.9198	.9197
11. Std. P.P.	.8584	.0847	.2023	.0595	.1358	.0131	.0054	.4498	1.0094	1.0092
12. V.C.	—.1098	.0459	—.0671	—.2287	.2128	.3580	—.2752	.2287	.3725	.3724
13. B.H. aft. Ex.	—.1643	.1346	—.2453	—.0765	.1587	.5595	—.3168	.1734	.5799	.5801
14. P.R. 2 min. aft. Ex.	.0263	.3569	.8508	—.2126	—.1475	.1002	.0461	—.2464	.9918	.9914
15. Diff: Std. P.R. & P.R. 2 min. aft. Ex.	.0836	.5975	.3834	—.1455	—.2697	—.0115	.1968	.0732	.6491	.6488
16. Diff: Hor. Std. P.R.	—.0875	—.3781	.1732	—.3767	.1591	.0457	—.3773	—.3089	.5885	.5882
17. Diff: Hor. Std. S.P.	—.1125	—.2484	.3630	.2351	—.1758	.2718	.4803	.5433	.8921	.8920
18. Diff: Hor. Std. D.P.	—.1803	—.2912	—.0584	—.5589	—.4315	.1703	.2675	.2370	.7761	.7757
19. Diff: Hor. Std. P.P.	.0681	—.0188	.3385	.7491	—.2961	.0406	.0640	.1489	.7963	.7961
TEST BATTERIES										
20. McCloy Test	—.4419	.0274	—.4439	.1043	—.3641	—.1234	—.0019	.6098	.9237	.9235
21. Barach Test	—.1159	—.2267	.8486	—.1207	—.0296	.0903	—.1264	.0744	.8301	.8298
22. Stone Test	.7490	—.0678	—.0412	—.0677	—.3174	.1808	—.2946	.1332	.8099	.8094
23. Basal Metab. Test	.6073	—.1021	.7572	—.0541	.1890	.1180	—.0391	.0191	1.0070	1.0066
24. Tigerstedt Test	.9091	.0702	.1060	.1086	.2338	.0660	.0888	.1999	.9606	.9611
25. P.P. x P.R. (Std)	.7877	—.0414	.5283	.0050	.1678	.0604	—.0013	.2427	.9920	.9917
26. P.P. x P.R. ÷ Dias. (Std.)	.8875	.0094	.3557	.0313	.2471	.0501	.0375	.0550	.9833	.9827
27. Crampton Test	—.0420	.0183	.2304	.3653	—.3733	.2558	.4519	.5613	.9126	.9126
28. Schneider Test	—.0852	—.1408	—.3863	.1733	—.2261	.0113	.2012	.6145	.6756	.6758
29. McCurdy-Larson Test	—.3429	.1684	—.1330	—.0804	.3141	.4856	—.1899	.3181	.6420	.6418

the factors. It will be noted (Table II) that the variables are not all completely described by the eight factors. This means that the tests which are not described are specific with respect to the variables in this experiment. Unless a factor is found in at least two tests, it cannot be isolated.

The process of factor rotation involved only the single variables and did not include the test batteries. The test batteries were included only for the purpose of determining their physiologic characteristics on the basis of the isolated factors.

Factor One.—Because of the high loadings (correlations) of sitting and standing pulse pressures (.76, .86) this factor is identified as *pulse pressure*. The multiple correlation of these two variables with the factor is .89. The horizontal pulse pressure correlates .69 and diastolic pressure (horizontal, sitting, and standing) also correlates significantly. The intercorrelations of diastolic pressure and pulse pressure, however, are not sufficiently high to predict one, given the other. Pulse pressure must therefore be considered as a single component when compared with the other cardiovascular-respiratory measures.

Factor Two.—The physiologic variables correlating significantly with this factor are: pulse rate two minutes after exercise (.36), the difference between standing pulse rate and pulse rate two minutes after exercise (.60), and the difference between horizontal and standing pulse rates (—.38). The multiple correlation of these variables is .65. This factor is therefore described as *pulse rate response to exercise in relation to normal.*

Factor Three.—The variables correlating significantly with factor three are horizontal pulse rate (.72), standing pulse rate (.83), and pulse rate two minutes after exercise (.85). The "net" contribution of the horizontal pulse rate in the presence of standing pulse rate and pulse rate two minutes after exercise is insignificant. The multiple correlation of the remaining two tests is .89. The factor is described as *pulse rate response to exercise.*

Factor Four.—The variables significantly related to the factor are horizontal diastolic pressure (.44), horizontal pulse pressure (—.57), difference between horizontal and standing diastolic pressures (—.56), and the difference between horizontal and standing pulse pressures (.75). On the basis of this evidence it is described as *pulse pressure in response to postural change.* The multiple correlation of these variables (eliminating horizontal pulse pressure) is .78.

Factor Five.—The physiologic variables correlating significantly are: sitting diastolic pressure (—.29), sitting pulse pressure (—.28), standing diastolic pressure (—.36), difference betwen horizontal and standing diastolic pressure (—.43), and the difference between horizontal and

standing pulse pressures (—.30). Sitting and standing diastolic pressures have the largest "net" significance in predicting the factor (.74). It is described as *diastolic pressure and changes with respect to body position.*

Factor Six.—Vital capacity and breath-holding 20 seconds after exercise correlate significantly (.36 and .56). These items yield a multiple correlation of .58 and are described as the *respiratory* factor. The multiple correlation is not significantly high to effectively predict the factor; however, the items are not significantly related to the other components; therefore, it can be considered another aspect of cardiovascular-respiratory function.

Factor Seven.—The items correlating significantly are horizontal systolic pressure (—.52), horizontal diastolic pressure (—.32), sitting systolic pressure (—:46), sitting pulse pressure (—.46), and difference

TABLE III
MULTIPLE CORRELATION OF VARIABLES WITH THE FACTORS

Factor	*Variable in the Factor**	*Multiple Correlations*
I	5. Horizontal Pulse Pressure	RI.5,8,11 = .894
Pulse Pressure	8. Sitting Pulse Pressure	
	11. Standing Pulse Pressure	RI.8,11 = .891
II		
Pulse Rate Response	14. Pulse Rate 2 Min. after Exer.	RII.14,15,16 = .654
to Exercise in Relation to Normal	15. Diff. in Std. Pulse Rate and Pulse Rate 2 Min. after Exerc.	
	16. Diff. in Hor. and Std. Pulse Rate	
III		
Pulse Rate Response	1. Hor. Pulse Rate	RIII.1,2,14 = .893
to Exercise	2. Std. Pulse Rate	
	14. Pulse Rate 2 Min. after Exerc.	RIII.2,14 = .892
IV		
Pulse Pressure in	4. Hor. Diastolic Pressure	
Response to Postural	18. Diff. in Hor. and Std. Diastolic Pressure	RIV.4,18,19 = .775
Change	19. Diff. in Hor. and Std. Pulse Pres.	
V	7. Sitting Diastolic Pressure	
Diastolic Pressure and	10. Standing Diastolic Pressure	RV.7,10,18,19 = .740
Changes with Respect	18. Diff. in Hor. and Std. Diast. Pres.	
to Body Position	19. Diff. in Hor. and Std. Pulse Pres.	RV.7,10 = .744
VI	12. Vital Capacity	
Respiratory Factor	13. Breath-Holding 20 Sec. after Exer.	RVI.12,13 = .577
VII	3. Hor. Systolic Pressure	
Blood Pressure Response	4. Hor. Diastolic Pressure	RVII.3,4,6,8,17 = .855
to Changes in Posture	6. Sitting Systolic Pressure	RVII.4,6,8,17 = .854
	8. Sitting Pulse Pressure	RVII.4,8,17 = .854
	17. Diff. in Hor. and Std. Systolic Pres.	
VIII	3. Hor. Systolic Pressure	RVIII.3,6,9,17 = .864
Systolic Pressure	6. Sit. Systolic Pressure	RVIII.6,9,17 = .864
	9. Std. Systolic Pressure	RVIII.9,17 = .862
	17. Diff. in Hor. and Std. Systolic Pressure	

*The Arabic number preceding the name of the variable corresponds to the numbers of the variable in Tables I and II.

between horizontal and standing systolic pressure (.48). The multiple correlation between horizontal diastolic pressure, sitting pulse pressure, difference between horizontal and standing systolic pressure, and the factor is .85. It is described as *blood pressure response to changes in posture*. The identification, however, is with a low degree of confidence because of the lack of sufficient degree of differentiation of the various pressures.

Factor Eight.—The variables correlating are all systolic pressures in the various positions (horizontal, sitting, and standing). Standing systolic pressure and the difference between the horizontal and standing systolic pressure correlate (multiple) .86. It is identified as *systolic pressure*.

INTERPRETATION OF CARDIOVASCULAR TEST "INDEX SCORES" IN TERMS OF THE IDENTIFIED FACTORS

One of the major purposes of this analysis is to determine the physiologic characteristics of the ten selected cardiovascular test "index scores." The "index scores" were therefore included in the factor analysis. This procedure does not establish the validity of these batteries as validity can be established only by its evaluation with respect to an acceptable criterion, but it simply yields a physiologic description on the basis of the identified factors. The procedure also assumes that the isolated factors resulting from the analysis of the single cardiovascular-respiratory variables are correctly identified. An example illustrates the use of the factor method in the analysis of the test batteries. A cardiovascular test battery is a complex phenomenon. We are interested in two characteristics of this battery, its validity and its physiologic characteristics. It is in the latter that the factor analysis procedure makes a contribution. If, for example, it is found that three factors correlate perfectly with the test battery, it can be assumed that the test battery is fully described by the three factors, and if the factors are correctly identified, the description of the test battery is established. So it is in this analysis. If diastolic pressure correlates perfectly with a test battery, the test battery is physiologically the same as diastolic pressure, assuming that no other variables are reacting on both the test battery and diastolic pressure to produce the perfect correlation. If a large portion of the relationship is due to age, it would be hardly correct to say that the correlation is perfect between the diastolic pressure and the test battery. For a complete physiologic description it is necessary to know what constitutes the relationship. If the variables could all be assumed to be pure, this problem would be eliminated, but this cannot be assumed to be true in circulatory-respiratory variables.

In Table IV the relationships between the eight isolated cardiovascular factors and the various test batteries have been listed in terms of significant correlations. It is noted that Factor One (Pulse Pressure) correlates significantly with five test batteries (Stone .75, Basal Metabolic

.61, Tigerstedt .91, P.P. x P.R. .79, and P.P. x P.R. ÷ Dias. .89). The reason for such correlations is apparent as the tests have for their major constituent pulse pressure. The descriptions by the other seven isolated factors are hardly significant, with the exception of the Basal Metabolic Test and the P.P. x P.R. which correlate .76 and .53, respectively, with Factor Three (pulse rate response to exercise).

TABLE IV

PHYSIOLOGIC DESCRIPTION OF CARDIOVASCULAR TESTS BY IDENTIFIED FACTORS

Identified Factors	*Significant Correlations with Tests and Identified Factors*	
	Test	*Correlation*
1. Pulse Pressure	1. Stone	.749
	2. Basal Metabolic	.607
	3. Tigerstedt	.909
	4. P.P. x P.R.	.788
	5. P.P. x P.R. ÷ Dias.	.888
2. Pulse Rate Response to Exercise in Relation to Normal		*No significant correlation*
3. Pulse Rate Response to Exercise	1. Barach	.849
	2. Basal Metabolic	.757
	3. P.P. x P.R.	.528
4. Pulse Pressure in Response to Postural Change	1. Crampton	.365
5. Diastolic Pressure and Changes with Respect to Body Position	1. McCloy	—.364
	2. Stone	—.317
	3. Crampton	—.373
	4. McCurdy-Larson	+.314
6. Respiratory Function	1. Crampton	.256
	2. McCurdy-Larson	.486
7. Blood Pressure Response to Changes in Position	1. Stone	—.295
	2. Crampton	.452
8. Systolic Pressure	1. McCloy	.610
	2. Crampton	.561
	3. Schneider	.615
	4. McCurdy-Larson	.318

Factor Two (pulse rate response to exercise in relation to normal) does not correlate significantly with any of the ten cardiovascular test batteries. The factor is included in some of the test batteries but the influence of this variable on the index score appears to be lost in the summing process. This is an important and comparatively unrelated factor.

The Barach Test (.85), the Basal Metabolic Test (.76) and the P.P. x P.R. Test (.53) all correlate significantly with factor three (pulse rate response to exercise). These tests all include this component as a major constituent. The factor is included in some of the other batteries (McCloy and Schneider), but the index scores of these tests do not reflect it.

The Crampton Test is the only test correlating with factor four (pulse pressure in response to postural change, .37). This test does not include the factor, but the slightly significant correlation is probably due to the relationship of the factor to systolic pressure postural change, which is a major element of the test. These two factors are related to approximately this degree.

Factor five (diastolic pressure and changes with respect to body position) correlates slightly with four tests (McCloy, Stone, Crampton, and the McCurdy-Larson Tests). The signs are not consistent in the McCurdy-Larson Test, but this may be due to the system used in scaling the diastolic pressure. The high and low diastolic values receive a low index score in the McCurdy-Larson Test. The factor is rotated as a negative factor, and hence the positive relationship with the McCurdy-Larson Test.

The respiratory factor (six) correlates significantly (.49) with the McCurdy-Larson Test and slightly with the Crampton Test. The McCurdy-Larson Test correlates because of vital capacity and breath-holding ability 20 seconds after exercise; the Crampton Test, because of the relationship of pulse rate to the factor. The correlations with the other tests are zero.

Factor seven (blood pressure response to changes in position) correlates significantly with the Crampton Test and slightly and negatively with the Stone Test. This is in accord with the nature of the factor and tests. The difference in the horizontal and the standing systolic pressures is a positive element of the factor and one of its significant elements. The Crampton Test includes this element, and also correlates positively with the factor. The horizontal systolic pressure has a negative correlation with the factor which indicates that, as the factor increases, the horizontal systolic blood pressure decreases. The sitting pulse pressure correlates with the factor, and this item is the major constituent of the Stone Test, and hence the negative correlation.

Factor eight (systolic pressure) describes partially the McCloy Test (.61), the Crampton Test (.56), the Schneider Test (.61), and the McCurdy-Larson Test (.32). The correlations are significantly higher in the first three tests. The Crampton and the Schneider Tests include this factor as a major component. In the McCloy Test, however, the only account which can be made is the relationship of systolic pressure to diastolic pressure in the standing position.

It will be noted in Table II that the eight isolated factors describe six of the ten test batteries from 90 percent to 100 percent,* two from 80 to 90 percent, and two from 60 percent to 70 percent. It is also inter-

*The factor loadings squared are interpreted as percentages.

esting to note that not one of the test batteries reflects all eight isolated factors. The tests correlating significantly with each of the eight isolated factors vary from one to five. The correlations, however, are not sufficiently high to predict the factor. It should be pointed out that the reason the tests may not reflect the various components may be because of the system of scaling. For example, diastolic pressure in the McCurdy-Larson Test is scaled according to the logical physiological interpretation; that is, the mean is used as the desirable value (best score) with the extremes as the undesirable values (low scores). The unscaled diastolic pressure will therefore not correlate with the scaled value, at least using the linear method. The index score, therefore, may have a higher degree of validity as an estimate of circulatory-respiratory function than is reflected by its correlation with the factor.

IMPLICATIONS OF THE FACTOR ANALYSIS FOR TEST CONSTRUCTION

The results of the factor analysis of the 29 variables and tests indicate two conclusions: (1) the selection of variables for the construction of the ten test batteries does not include all the factors describing circulatory-respiratory function, and (2) the "index scores" of the ten selected tests do not reflect the isolated factors of circulatory-respiratory function. It was also found that in some test batteries, where the factor represents a major part of the battery, the "index score" does not reflect it to a significantly high degree. This is because of the influence of the other items in the battery. It does not necessarily mean that these tests are invalid with respect to a specific objective, but it does mean that they are invalid as a complete index (in terms of the selected variables) of circulatory-respiratory status.

The implications of these findings for test construction are three: (1) the selection of the items for the complete estimate of cardiovascular-respiratory function should include all factors of those systems; (2) if some of the factors are to be eliminated, this must be done in terms of a specific criterion, which represents the use to be made of the selected factors; and (3) the sum of a number of cardiovascular-respiratory measures may be a better index of a specific ability, but it does not reflect the factors which are constituent parts. With respect to the individual factors, the index score has no diagnostic value. In order to know the functional status of the individual each factor must be considered individually. If one is interested in knowing the circulatory respiratory status of the individual, all factors must be considered, with the sum of the scores, if desired, as yielding *additional* information. If one is interested in describing a particular ability, the items should be selected which significantly describe it with a "weighted" sum prepared (regression equation).

CONCLUSIONS

1. A factor analysis of some selected cardiovascular-respiratory variables shows eight factors. These have been identified as pulse pressure, pulse-rate response to exercise in relation to normal, pulse-rate response to exercise, pulse pressure in response to postural change, diastolic pressure and changes with respect to body position, respiratory function, blood pressure response to changes in position, and systolic pressure. Other factor analysis studies (not comparable with respect to specific factors) show a minimum of eight factors describing approximately the same selection of variables. N = 145

2. Utilizing the results of the factor analysis and in addition a logical analysis of the table of intercorrelations, the following variables were selected as being sufficiently unrelated as to be considered independent:

a. Horizontal pulse rate.
b. Difference between horizontal and standing pulse rates.
c. Difference between normal standing pulse rate and pulse rate two minutes after exercise.
d. Horizontal systolic pressure.
e. Difference between horizontal and standing systolic pressures.
f. Horizontal diastolic pressure.
g. Standing diastolic pressure.
h. Difference between horizontal and standing diastolic pressures.
i. Horizontal pulse pressure.
j. Difference between horizontal and standing pulse pressures.
k. Breath-holding 20 seconds after exercise (as found in the McCurdy-Larson Test).
l. Vital capacity.

3. The implications of the factor analysis of the cardiovascular-respiratory variables and tests for test construction are: (a) the selection of the items for a complete evaluation should include all factors; (b) if some of the factors are to be eliminated, this must be done in terms of a specific criterion, which represents the use to be made of the selected factors; and (c) the sum of a number of variables may be a better index of a specific ability, but it does not reflect the factors which are constituent parts.

4. The eight factors isolated in this research are not described by any one of the ten test "index scores." The reasons for the lack of relationship are: (1) the factor may not be included in the battery; (2) the number of items in the battery may be so numerous that the influence of the factor is not reflected; (3) the system of scaling the items in the test batteries makes the same items incomparable; and (4) the test battery may yield information not included in the eight factors.

BIBLIOGRAPHY

1. Larson, L. A. *A Factor Analysis of Some Cardiovascular-Respiratory Variables With Particular Reference to the Schneider and the McCurdy-Larson Tests.* (Restricted). Civil Aero. Adm. Div. of Res. Report No. 17, Washington, D. C.

2. ....., "A Study of the Validity of Some Cardiovascular Tests," *Journal of Exp. Ed.*, 7:3 (March, 1939), pp. 214-220.

3. McCloy, C. H., "A Study of Cardiovascular Variables by the Method of Factor Analysis," *Proceedings, Second Biennial Meeting*, Soc. for Res. In Child Development, Oct., 1936, pp. 107-113.

4. McCurdy, J. H., and L. A. Larson. *The Physiology of Exercise.* Philadelphia: Lea and Febiger, 1939, pp. 315-336.

5. ....., "The Validity of Circulatory-Respiratory Measures As An Index of Endurance Condition in Swimming," *Research Quarterly.* 11:3 (Oct., 1940), pp. 1-11.

6. ....., "Age and Organic Efficiency," *The Military Surgeon*, 85-2 (Aug., 1939), pp. 93-103.

7. Murphy, Mary A., "A Study of the Primary Components of Cardiovascular Tests," *Research Quarterly*, 11:1 (March, 1940), pp. 57-71.

8. Quamme, H. J. "A Factor Analysis of Cardiovascular Variables," Unpublished master's thesis, Springfield College, Springfield, Mass.

9. Thurstone, L. L., "Current Issues in Factor Analysis," *Psychol. Bull.*, 1940, pp. 37, 189-236.

10. ..... *A Simplified Multiple Factor Method.* Chicago: The University of Chicago Bookstore, 1933.

11. ..... *The Vectors of the Mind.* Chicago: The University of Chicago Press, 1935.

The Relationships of External Fat To Physical Education Activities and Fitness Tests*

By RAMON W. KIREILIS
AND THOMAS K. CURETON
University of Illinois Physical Fitness Laboratory
Urbana, Illinois

(Submitted for publication, December, 1946)

THE PROBLEM

THIS study aims to show the relationship between fat as measured by calipers at six places on the exterior of the body and several aspects of physical education work:

1. Relationships to certain structural physical tests, i.e., ponderal index, calf girth, vital capacity residuals, biceps girth, and somatotype rating.

2. Relationships to certain functional organic efficiency tests, i.e., strength, strength per pound of body weight, expiratory force, flarimeter breath holding after chair stepping for 60″, and various pulse rate and blood pressure tests.

3. Relationships to motor fitness performances on the Illinois 14-Item Motor Fitness Test, the Cureton 18-Item Motor Fitness Test, and the Mile Run.

I. SURVEY OF RELATED STUDIES

A. *Trends of Relationship Between Health and Fat on the Body.*— The importance of obesity is amplified by the Metropolitan Life Insurance Co. upon the issuance of the following statement:

> Men who are 35 percent or more above average weight for their age and height have a mortality rate of 1½ times that of the average man; also, when the abdominal girth is more than 2 inches greater than the chest girth at full expansion, the extra mortality rate is 50 percent above the excess mortality associated with the overweight itself. Diseases such as high blood pressure, albuminaria, glycocuria, cerebral hemorrhages, angina pectoris, and Bright's disease show unusually high indices with the obese individuals.[22]**

*All of these tests are described and standard score rating tables are given in Cureton's *Physical Fitness Workbook*, 1944 edition or 1947 edition published by C. V. Mosby Co., St. Louis.

**All reference numbers hereafter used refer to the alphabetically arranged bibliography at the end of the paper.

Obese individuals usually have a high frequency of the following diseases: hypertension,[2] functional disease of the heart, kidneys and pancreas (diabetes),[16] gallstones, heart lesions, coronary artery disease, lung disease, arthritis and exhausting bronchitis.[12]

Obese individuals not eating normal amounts of carbohydrates and eating abnormal amounts of fats and proteins are liable to a condition known as acidosis. These obese individuals cannot participate in long continuous physical activities because of the deficiency in carbohydrates and the excess of fat stored in the depots. This condition of acidosis causes headaches, nausea, and occasionally vomiting, weakness, sleeplessness, muscular aches, acid mouth, sour stomach, acid urine, acid sweat, "sour disposition" and blood vessel disease which leads to other complications.[24]

B. *Relationships Between Adipose Tissue and Motor Performances.* The obese individual is inefficient in work production that is long and continuous. It is no wonder that an obese individual consuming a high fat diet is unable to participate in long continuous endurance activities.[18] Excessive fat results in the loss of "tonus" of the abdominal wall which affects the blood supply, circulation, and the elimination of waste products[19] and causes dilation of the heart and poor circulation of the blood.[9] The vital organs become surrounded with fat and are weakened.[21] The excess fat interferes with breathing, and often results in fallen arches, poor posture, a bone disease, and an increase in weight.[22] Sweating is more profuse due to the insulating effects of adipose tissue. The obese individual becomes overheated.[16]

Every pound of fat requires 4,500 feet of blood vessels.[10] The obese person is like an overloaded truck, usually in danger.[26] A great deal of the blood goes to fat instead of the nerves, muscles, and vital organs.[9] The added weight in muscular exertion places a greater load upon the heart and circulatory system.[2]

The obese individual is usually awkward,[28] and the muscles become flabby from the lack of exercise.[26] He is usually not agile and loses free movement in limbs and joints.[18] He is seldom efficient. Weight is a reliable index of the state of a person's physical fitness.[26] The obese individual has a high frequency of accidents and minor sprains.[28]

Few studies have been made showing the direct relationship between specific motor performances and adipose tissue. The use of calipers to measure the fat overcoat is a relatively recent development in physical fitness measurement. Grover's[16] findings (1936) were of some significance. The correlations he obtained between motor ability and adipose tissue are summarized as follows: The Rogers' Physical Fitness Index correlated —.404 ± .046 with adipose tissue and indicates that the physically fit individuals have less adipose tissue, based upon Rogers' normative tables. The Cozens' All-around Athletic Ability Test correlated —.202 ± .504 with the Grover-Cureton fat index. Slightly

better performance in athletics is obtained when less fat is on the body; other correlations obtained showed negligible relationships. These are: 440-yard run correlated with adipose tissue —.113 ± .067; McCurdy-Larson Organic Efficiency Test correlated with adipose tissue —.056 ± .058; and 100-yard swimming endurance correlated —.064 ± .094 with adipose tissue, using Cureton's 5-lap Drop-off Endurance Test.

Studies made by Cureton (1941)[6] show that endomorphs and ectomorphs gave poor performances in comparison with mesomorphs in the Rogers' Strength Index, sprinting, 440-yard run, broad jump, flexibility, Brace Test, and the Cozens' Test. Hall (1942) found obese farm boys to be relatively weaker in leg-lifts on the high bar, chins, push-ups, and jumps.

Bruch (1940-41)[4] found that the majority of obese individuals were inactive. Boothby (1922)[3] found that the metabolic rates of 98 percent of the obese men were within the range of + 15 and — 15; however, Lucas (1933)[19] found that the obese individuals had a lower metabolic rate than the linear type of individual within this range. These authors suggest the great importance of the lack of exercise and a fattening diet as the principal causal factors as only 2-4 percent of all obese subjects can attribute the cause to metabolic disorder rather than to habits of overeating and inadequate exercise.

Steinhaus has attempted to show the public how difficult it is to lose a pound of fat through exercise. Listed below are the tables of "human mileage." The tables show the amount of exercise or work that is needed over and above that required for just lying still. These tables are approximate and assume that the individual is of average size and weighs 155 pounds. This type of an individual, in order to burn off a pound of fat, must perform the following amounts of work. The following are direct quotations from Steinhaus[25]:

1. Run 129 one hundred-yard dashes at ten seconds each.
2. Run 17.3 miles at a rate of a mile in six minutes.
3. Play strenuous football for four hours and forty-eight minutes.
4. Wrestle five and one-half hours.
5. Use an electrical vibrator machine for 375 fifteen-minute periods.
6. Walk thirty-four and three-fourths' miles, at a rate of five miles per hour.
7. Fence for eight hours.
8. Saw wood for ten and one-half hours.
9. Do parallel bar work for eight and one-half hours.
10. Play ping-pong for seventen and one-fourth hours.
11. Walk sixty-six and one-half miles, at a rate of mile in seventeen and one-half minutes.
12. Drill for twenty and one-half hours.
13. Walk 144 miles at a rate of two miles per hour.

14. Do ordinary office work for eleven eight-hour days.
15. Shovel 114,739 pounds of sand into a wheelbarrow.
16. If he is a skilled mason, lay 14,731 bricks.
17. Walk to the top of the Washington Monument forty-eight times.
18. Do 5,714 push-ups from the floor.

Generally, Steinhaus claims that it is unsafe for obese individuals to attempt to reduce through exercise. The appetite is stimulated through the participation in exercises, and the loss of weight is mainly due to a loss of water, which is regained later. He recommends that the best means to lose fat is through diet.

C. *Experimental Effect of Exercise Upon Adipose Tissue as Measured by Calipers.*—The utilization of fat through exercise has been shown by Anderson and Lusk (1917),[1] Rapport (1929),[23] and Gemmill (1942).[15] Also Gardiner (1925),[14] and Courtice and Douglas (1929)[5] have shown that the utilization of fat is higher when the intensity of the exercise increases. Courtice and Douglas (1929) and Dill (1935)[11] found that the burning of carbohydrates is more efficient than protein and fat during exercise. The studies of Hotzbaur (1927)[17] and McCurdy (1939)[20] indicate that weight can be lost through exercise. However Tuttle (1943)[27] advocates diet and exercise. Tuttle found that loss of up to 5 percent of the body weight does not have any detrimental effect on the body.

No fat was actually measured in these studies. No other studies were found in searching the literature which measured the reduction of adipose tissue concomitantly with the practice of specific exercises.

II. MEASUREMENT AND STATISTICAL RELATIONSHIPS OF FAT TO PHYSICAL PERFORMANCES

At the University of Illinois measurements of external fat were made on students in various physical education classes and these same students were given a large number of physical fitness tests. The results were correlated by the Pearon scattergram statistical procedure to determine the relationships of fat to these performances.

Techniques of Measurement for External Fat.—The Franzen type calipers[13] were used for measurement of the external fat. The calipers used had 200 gms. of force at 20 mm. and 400 gms. of force at 30 mm. of opening between the jaws. In making the measurements, the calipers were held in the right hand. The volar surface of thumb and index finger of the left hand pinched up the adipose tissue to free it from the muscles. Then the measurements were made by placing the jaws of the caliper blades below, parallel and in line with the tips of the thumb and forefinger. The pressure of the spring of the calipers held the tissue firmly

and the indicator on the scale gave a measure of the thickness of the adipose tissue in millimeters. When the tissues were very firm and the calipers began to slip, the blades of the calipers were then pressed inward just enough to keep the tips of the blades in line with the left thumb and forefinger.[13, 7] Therefore, the absolute measurements include a double layer of adipose tissue and skin. The skin thickness was neglected.

The specific techniques of measuring adipose tissue were as follows. The technique used was identical with the technique used by George H Grover.[16*]

Cheeks: Measurment was made on the right cheek, level with the lower part of the nose, at the center of the cheek.

Abdomen: The measurement was made two inches to the left of the navel.

Hips: Measurement was made on the right hip at a point located at the center of the sagittal plane of the body, and even with the top of the crest of the illium.

Gluteal: Measurement was made on the left buttocks, level and a little above the gluteal fold.

Rear Thigh: Measurement was made at the center of the rear thigh, half-way between the knee and cleft of the gluteal fold with the foot relaxed and knee bent.

Front Thigh: Measurement was made at the center of the front thigh half-way between the knee and the groin with the foot relaxed and off the ground.

Cureton (1941) made a normative study on adipose tissue of young men using the Grover-Cureton Weighted Adipose Tissue Index, developed at Springfield College, on young athletic subjects. Rating tables for total weighted adipose tissue are provided but in a further study at Illinois University (1944) tables were provided for fat on the cheeks, abdomen, hips, front thigh, rear thigh, and the total based on direct measurement of external fat of 332 cases of young men at the University of Illinois.

Reliabilities of the Measurements.—Grover and Cureton have found high reliabilities for the adipose tissue measurements. The reliabilities for the six measurements are as follows:

Cheeks	.862 ± .028	Gluteals	.968 ± .006
Abdomen	.975 ± .005	Front Thigh	.968 ± .006
Hips	.901 ± .019	Rear Thigh	.944 ± .001

The reliability for the six measurements summed was found by Cureton to be .90. Grover did not give the composite reliability. These reliabilities as a whole are satisfactory.

A. *Statistical Findings, 1942 Data Showing Relationships to Various Physical Tests.*—Correlations were computed between the Adipose

*Cf. Cureton, *Physical Fitness Workbook*, p. 110, for illustration.

Tissue Index (Sum of 6 measurements on cheeks, abdomen, waist, front thigh, rear thigh, gluteal fold in mm.) and 23 various physical performances, using 113 young male subjects in the basic physical fitness classes at the University of Illinois. The results are as follows:

TABLE I

CORRELATIONS BETWEEN ADIPOSE TISSUE AND MOTOR PERFORMANCE
(Young Men, Basic Physical Fitness Course, PEM 60, University of Illinois)

Motor Skills	r	r* (true)	P.E.	r_{II}
Composite Running Index**	—.492	—.578	±.047	+.80
Mile Run	—.399	—.458	±.033	+.84
Chins	—.397	—.427	±.032	+.95
1000-Yard Run	—.398	—.436	±.037	+.86
100-Yard Dash	—.337	—.399	±.036	+.79
2-Mile Run	—.336	—.382	±.036	+.80
Broad Jump	—.326	—.360	±.038	+.90
3½-Mile Run	—.287	—.331	±.042	+.80
Fence Vault	—.286	—.307	±.048	+.95
Stick Body	—.210	—.264	±.045	+.71
Frog Stand	—.167		±.041	+.92
Hops	—.146		±.040	+.82
Dive and Roll	—.117		±.036	+.95
V-Sit	—.112		±.038	+.80
Sit-Ups	—.072		±.038	+.83
Trunk Extension Flexibility	—.066		±.058	+.94
Man Lift	—.066		±.036	+.87
Leg Lifts	—.056		±.037	+.73
Medicine Ball Put	—.009		±.037	+.82
Diver's Stance	—.004		±.048	
Schneider Index	—.010		±.045	+.86
Trunk Flexion	—.044		±.061	+.95
Extension Press-Ups	—.053		±.053	+.95

*Corrected for attenuation. Cf. H. E. Garrett, p. 213, *Statistics in Psychology and Education*, 1933:

$$\sqrt{AB} = \frac{\sqrt{A_1B_1}}{\sqrt{A_1A_2}\ \sqrt{B_1B_2}}$$

**Composed of the sum of the standard scores for mile, 2-mile, 1000-yard drop-off and 3½-mile steeplechase runs. The other events are described in previously published sources.[7]

B. *Statistical Findings, 1945 Data.*—Correlations of the Adipose Tissue Index with 23 other physical tests and measurements gives evidence to show how external fat is related to various aspects of physical fitness, including physique, organic reactions to exercise, and motor performances. Cureton obtained the following correlations on 110 University of Illinois freshmen in the medical school (18 to 25 years of age):

TABLE II

CORRELATIONS BETWEEN ADIPOSE TISSUE AND VARIOUS PHYSICAL FITNESS TESTS

(N = 113)

	r	P.E.
Adipose Index with the Ponderal Index	+.676	±.035
Adipose Index with Calf Girth	+.627	±.037
Adipose Index with Vital Capacity Residuals	—.525	±.047
Adipose Index with Biceps Girth	+.542	±.046
Adipose Index with Strength/lbs. of Weight	—.406	±.053
Adipose Index with Skeletal Index	+.399	±.054
Adipose Index with Illinois 14-Item Motor Fitness Test	—.345	±.056
Adipose Index with Composite of the Standard Scores for the Mile Run, 14 and 18-Item Motor Fitness Tests	—.340	±.056
Adipose Index with Criterion of 22 Tests (Sum of Standard Scores)	—.218	±.060
Adipose Index with Expiratory Force on U-Tube	—.203	±.061
Adipose Index with Mile Run	—.202	±.061

The results in Table III show fairly high negative correlations between the external fat as measured on six different places on the body and the composite endurance running criterion composed of the sum of the four standard scores for the mile, 2-mile, 3½-mile steeplechase, and the 1000-yard drop-off index. This means that less fat goes with better running.

TABLE III

CORRELATIONS BETWEEN COMPOSITE RUNNING INDEX AND THE FAT OF THE CHEEKS, ABDOMEN, HIPS, GLUTEAL FOLD, REAR THIGH, AND FRONT THIGH

Location of Measurement	r	r (true)	P.E.
Abdomen	—.640	—.737	±.037
Gluteal	—.498	—.588	±.049
Front Thigh	—.394	—.424	±.059
Hips	—.381	—.409	±.059
Rear Thigh	—.374	—.425	±.059
Cheeks	—.262	—.262	±.060

III. EFFECT OF STRENUOUS TREADMILL RUNNING ON EXTERNAL FAT

Three relatively fat students were selected from a class in physical education at the University of Illinois in the summer of 1944. They were of the endo-mesomorphic type and are shown in Illustration I. The body weight and the six adipose measurements were made at the start of the experimental work and at the start and end of every day of the experiment thereafter. Measurements were also made at the start of each period of the Schneider Index. The subjects ran several times each period, doing all they could do on each trial.

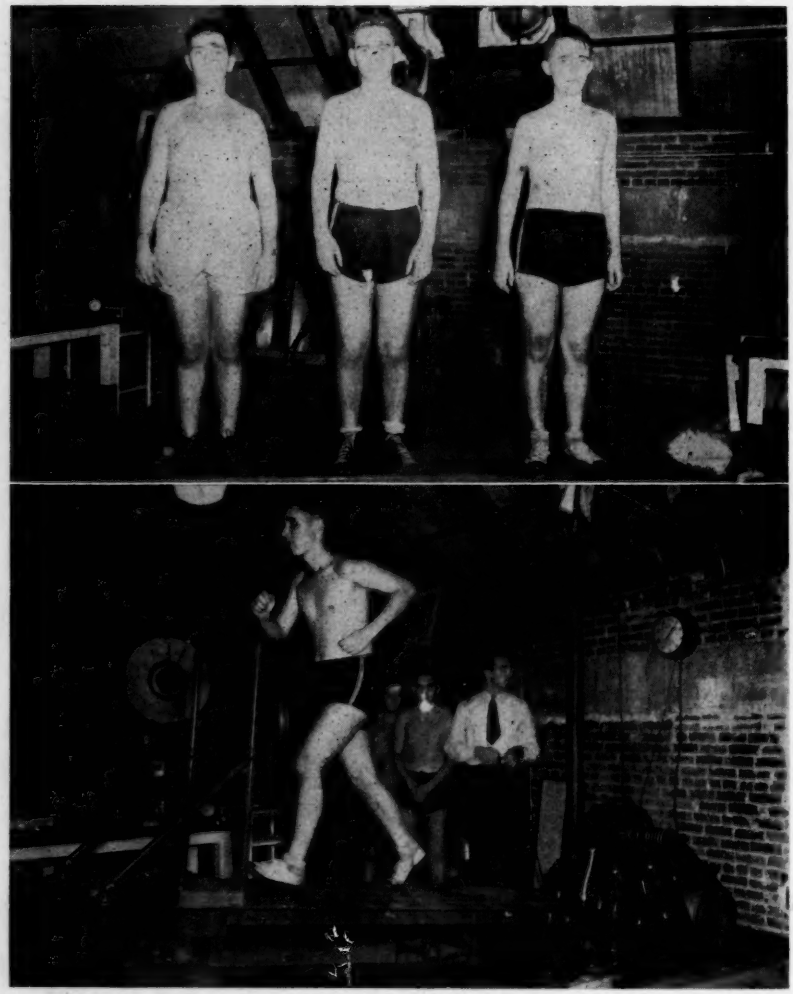

Illustration I. Top: The experimental subjects; Bottom: The method of exercising on the treadmill.

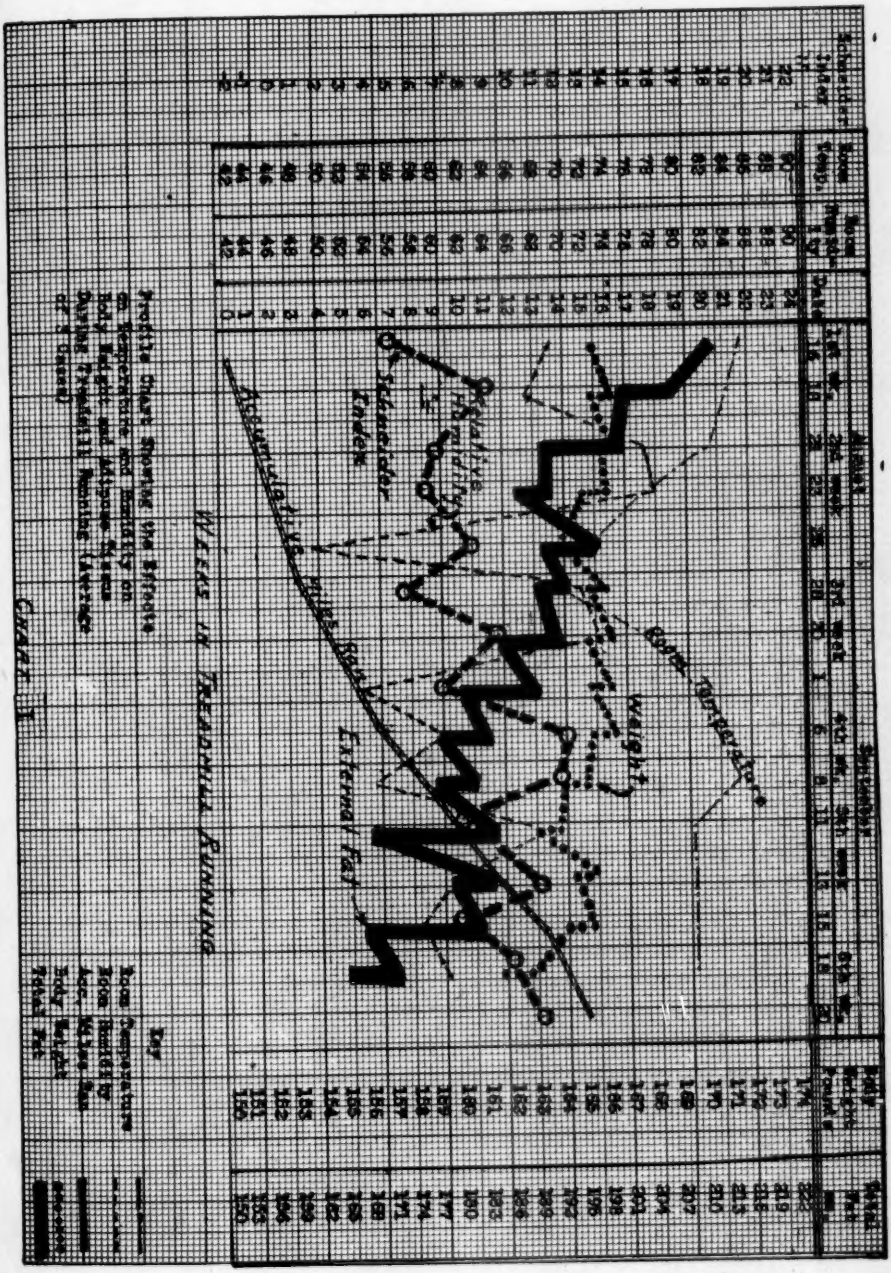
Chart 1
Weeks in Freshman Running
External Fat
Body Weight
Body Temperature

Chart I shows the results of six weeks of strenuous running. There is a definite trend for the external fat to diminish as the miles run accumulated from week to week. There was not much weight loss, which suggests that the men added muscle density as the fat reduced. The temperature and relative humidity were recorded each period but were not controlled as there were not facilities for such atmospheric control.

The fat loss at each place measured on the body over the whole six-week period is shown in Table IV.

TABLE IV.

EXTERNAL FAT LOSSES AT EACH PLACE ON THE BODY
(N = 3)

Location	*Start*	*End*	*Differences*	*Per Cent Differences (mm.)*	C.R.
Gluteals	37	27	10	27.1%	1.228
Rear Thigh	39	28	11	28.2%	1.307
Front Thigh	39	34	5	7.8%	.742
Cheeks	25	22	3	12.0%	.867
Abdomen	37	31	6	16.2%	.742
Hips	37	25	12	32.4%	1.462

It appears that relatively more fat was lost from the hips, rear thigh, and buttocks in this up-hill running than from the cheeks, front thigh, and abdomen. However, since the critical ratios for the differences between the means are all lower than 3, these differences cannot be guaranteed as to statistical significance.

CONCLUSIONS

1. Previous studies on body fat have not related measurements of external fat by the caliper method to various amounts of exercise or types of class work in physical education.

2. There are significant negative correlations of the magnitude of —.578 to —.264 between performances of strenuous physical exercises and external fat on the body, the correlations with endurance running being relatively greater than for other test exercises. Fat is a real handicap in most strenuous exercises.

3. Massive bodies (by the ponderal index) tend to have greater amounts of external fat and the larger calves and biceps are usually larger because there is more external fat present.

4. Cardiovascular and respiratory measurements do not seem significantly related to external fat within the range of the group data considering the low correlations obtained (—.203 to +.122) but in the fat reduction experiment using three subjects in treadmill running there was a progressive gain in the Schneider Index as the training proceeded.

5. In strenuous endurance running the fat on the abdomen and buttocks seems to be more of a handicap (with negative correlations with performances, —.737 to —.588) than fat on the thighs and cheeks (—.394 to —.226).

6. Strenuous treadmill running in August and September caused an appreciable loss of external fat over six weeks of training in 30-minute periods, three times per week. The loss was proportional to the miles cumulatively run, considered as the total distance added for a series of short runs in each period, summed over all the periods and converted to miles.

8. Weight is not a good guide to the fat loss, possibly because fat loss is compensated for by increased muscular density due to the exercise. The fat loss was at a greater rate than the weight loss in this experiment.

9. The average Schneider Test scores began at 5 for the first period and finished at 11 on the last period of the sixth week of training, a gain of 120 percent from the starting level in raw scores, equivalent to a gain from 30 to 50 on the standard score tables,* a standard score gain of 20 percent on the 100-point scale based on a normal distribution. In all but the 4th week of training the Schneider Index was regularly lower on Monday mornings after the week-end and improved during the week with treadmill running every other day.

BIBLIOGRAPHY

1. Anderson, R. J., and C. Lusk, "Animal Calorimetry: The Interrelation Between Diet and Body Condition and the Energy Production During Mechanical Work," *Journal of Biological Chemistry*, 32: 421-445, 1947.

2. Best, C. H., and H. B. Taylor, "Metabolism and Starvation-Obesity and Underweight," *Physiological Basis of Medical Practice*. Baltimore: Williams and Wilkins, 1943, pp. 1021-1028.

3. Boothby, W. B., and Irene Sandiford, "Summary of the Basal Metabolism Data on 8,614 Subjects With Especial Reference to the Normal Standards for the Estimation of the Basal Metabolism Rate," *Journal of Biological Chemistry*, 54: 783-803, September, December, 1922.

4. Bruch, H., "Energy Expenditure of Obese Children," *American Journal of Diseases of Children*, 60:1082-1109, 1940.

5. Courtice, F. C., and C. G. Douglas, "The Effects of Prolonged Muscular Exercise on the Metabolism," *Proceedings of the Royal Society London*, Series B., 119:381-439, (March 2, 1936).

6. Cureton, T. K., "Body Build as a Framework of Reference for Interpreting Physical Fitness and Athletic Performance," *Research Quarterly*, 12:2, 301-330, Supplement, May, 1941.

7. ————————. *Physical Fitness Workbook*. 2nd edition. Champaign, Illinois: Stipes Publishing Co., 1944, 179 pp.

8. ————————. *Physical Fitness Workbook*. 1st edition. Champaign, Illinois: Stipees Publishing Co., 1942, 150 pp.

*Physical Fitness Workbook, p. 135.

9. ..... *Basic Principles of Physical Fitness.* Minneapolis: Imperial Printing Co., 4 pp.

10. Deaver, G. G., "Obesity and Its Treatment," *Journal of Physical Education,* 36:63, March-April, 1939.

11. Dill, D. B., "The Energy Requirements in Strenuous Muscular Exercise," *The New England Journal of Medicine,* 213:532-534, September 12, 1935.

12. Forris, A. W., "Reduction of Obesity," *Medical Clinics of North America,* 8:336, July, 1924.

13. Franzen, Raymond. *Physical Measures of Growth and Nutrition.* School Health Research Monographs, No. II. New York: American Child Health Association, 1927, pp. 138.

14. Gardiner, Hill H., "Modern Technique in Treatment, The Treatment of Obesity," *Lancet,* 209:1034-1035, November 15, 1925.

15. Gemmill, C. L., "The Fuel for Muscular Exercise," *Physiological Review,* 22:32-53, January, 1942.

16. Grover, G. H. *An Approximate Analysis of Adipose Tissue in Relationship to Nutrition, Condition, and Skeletal Size.* Springfield, Mass.: unpublished master's thesis, Springfield College, 1936.

17. Hotzbaur, Ida C., "Reducing the Overweights," *American Physical Education Review,* 32:297-299, April, 1927.

18. Kahn, F. *Man in Structure and Function.* New York: Alfred A. Knopf, 1943, pp. 47-48.

19. Lucas, W. B., and H. B. Pryor, "The Body Factor Build Factor in the Basal Metabolism of Children," *American Journal of Diseases of Children,* Part I, 46:941-948, 1933.

20. McCurdy, J. H., and L. A. Larson. *The Physiology of Exercise,* Philadelphia: Lea and Febiger, 1939.

21. McLester, J. S., "The Principles Involved in the Treatment of Obesity," *Journal of the American Medical Association,* 81:2103-2105, No. 26, June 28, 1924.

22. Metropolitan Life Insurance Company, *Statistical Bulletin,* 18:5, May, 1937.

23. Rapport, D., "The Utilization of the 'Waste Heat' of Metabolism in Muscular Exercises," *American Journal of Physiology,* 91:238-253, December-January, 1929-30.

24. Sansom, W. D. *The Normal Diet.* 3rd ed., revised. St. Louis: C. V. Mosby Co., 1930, pp. 28-49

25 Steinhaus, A. H, "Why Exercise?" *Journal of Health and Physical Education,* 5:5, 5-7, May, 1934.

26. Styles, W. A. L., "The Campaign Against Corpulence," *Hygeia,* 12:599-602, July, 1934.

27. Tuttle, W. W., "The Effect of Weight Loss by Dehydration and the Withholding of Food on the Physiological Responses of Wrestlers," *Research Quarterly,* 14:2, 156-164, May, 1943.

28. Veeder, B. S., "The Overweight Child," *Journal of the American Medical Association,* 83:486-489, No. 7, August 16, 1924.

The Sexual Maturing of Girls As Related To Growth in Strength

By HAROLD E. JONES
University of California
Berkeley

(Submitted for publication, December, 1946)

THE association of sexual maturing with specific changes in growth rates has been clearly shown by Shuttleworth[8*] in a study of the physical growth of girls classified according to their age at menarche. A similar relationship between sexual maturing and functional traits (e.g., static dynamometric strength) may be less easy to demonstrate because differential factors associated with incentive, fitness, and physical activity may tend to obscure basic growth phenomena. If, however, individual differences in physical abilities are primarily determined, in a normal sample, by intrinsic rather than extrinsic factors, we would expect growth curves in different menarcheal groups to reflect differences similar to those already demonstrated for physical structures.

Prior to the analysis of age changes in strength, it may be useful to compare strength data for pre-menarcheal and post-menarcheal girls of the same chronological ages. This method has been employed, in cross-sectional studies, to examine the influence of physiological maturing upon various characteristics. It has not always rested upon adequately reliable data, for the age at menarche is sometimes based on retrospective reports, obtained from several months to several years after the event. In the present investigation longitudinal records have been used, collected in connection with a study of adolescent growth. Table I presents a comparison of the average strength of grip** for groups known to be pre- or post-menarcheal. At each of the three age levels marked differences occur; these differences are significant at the .1 percent level of confidence.

Table II presents similar data, in somewhat different age classifications, for a "Total Strength" composite, consisting of Right Grip, Left Grip, Pulling Strength, and Thrusting Strength.*** At ages 12.25 to 13.25, the differences between the two groups are about 1 S.D.; more than 80 percent of the post-menarcheal girls excel the mean for the less mature girls of the same age. After 13.25 the difference becomes smaller,

---

The material in this report is presented in greater detail as a chapter in a forthcoming book, *Motor Performance and Growth.* [6]

*Superior figures refer to numbered bibliography at end of article.

**Strength of Grip was tested for right and left hand separately, using the best of three records for each hand.

***The technique of measurement is presented in reference No. 6. The data of Table II are adapted from a master's thesis by Caldwell[5] based on the girls of the present sample. The class interval in Caldwell's study is from 12.0 to 12-49 years, etc.

as a result of continued growth among the less mature and the reduction or cessation of growth among those well past the menarche.

TABLE I

COMPARATIVE STANDARD SCORES* (RIGHT GRIP) FOR GIRLS OF DIFFERING PHYSIOLOGICAL STATUS

	Pre-menarcheal			*Post-menarcheal*			
*Age***	*N*	*Mean (kg.)*	*S.D.*	*N*	*Mean (kg.)*	*S.D.*	*t*
12.5	56	44.5	9.4	23	54.1	8.4	4.2
13.0	46	43.8	10.0	38	50.7	8.7	3.3
13.5	26	38.7	8.9	50	49.0	8.2	5.0

*Raw scores at each age have been transformed into scale scores, relative to a mean of 50 for the total group at that age, and with an S.D. of 10.

**The class interval is 12.25 - 12.74, etc.

TABLE II

A COMPARISON OF PRE- AND POST-MENARCHEAL CASES IN TOTAL STRENGTH

	Pre-menarcheal			*Post-menarcheal*			
Age	*N*	*Mean (kg.)*	*S.D.*	*N*	*Mean (kg.)*	*S.D.*	*t*
12.25	63	86.0	16.30	15	100.6	15.55	3.11
12.75	44	92.0	16.35	24	105.6	15.30	3.31
13.25	25	93.0	16.20	37	108.8	16.35	3.68
13.75	20	97.1	16.35	54	105.6	16.50	2.00
14.25	17	100.7	19.50	63	107.9	18.00	1.42

Tables I and II demonstrate that among girls of the same chronological age, those who are past the menarche are stronger than those who have not reached the menarche. We may interpret this as due to the fact that the former group have had the full effect of the pre-menarcheal growth spurt.

The anatomical basis for the relationship between strength and sexual maturity is to be seen in Reynolds'[7] recent report concerning the comparative growth of tissues as measured from x-rays; in the growth curves which he shows for early- and late-maturing girls (classified by breast development), curves for breadth of muscle begin to diverge, for the two groups, as early as 8 years. In the prepuberal years the early- and late-maturing appear to be more sharply differentiated by growth of muscle than by growth of bone or fatty tissue.

As a further approach to a study of the factors represented in Tables I and II, correlations were determined between Strength of Grip and (1) chronological age and (2) age deviation from menarche. These correlations were based on successive testings, at half-year intervals, of

the children at three school grade levels. The results, in terms of Pearson *r*, are given in Table III.

TABLE III

THE CORRELATION OF STRENGTH WITH CHRONOLOGICAL AND PHYSIOLOGICAL AGE

Average Age	N	*Strength of Right Grip Correlated with C. A.*	N	*Strength of Grip Correlated with Age Deviation from Menarche*
11.5	92	.196	79	.429
12.0	104	.201	90	.385
12.5	99	.193	90	.417

Within the restricted age range of each grade the table shows a moderate correlation (significant at the .1 percent level) between strength and maturity expressed in terms of the number of months from menarche.* A consistently lower correlation occurs between strength and chronological age. At any one age the difference between coefficients is significant at only the 5 percent level, but the stability of results in three successive testings deserves to be noted. It should not be surprising to find that strength depends more on individual maturity than on the calendar of years from birth; evidence of this nature, however, fails to show the precise temporal relationships between sexual maturing and the growth curve for strength.

At this point, we may utilize the more characteristic advantages of longitudinal data by examining age changes in terms of cumulative growth curves. Three groups of girls were selected from the Adolescent Growth Study [4, 5] with reference to their age at menarche. Representing the two extremes and the middle of the normal sample, these may be considered to be early-maturing, average, and late-maturing. The early group included only cases in which the menarche occurred before the age of 12 years; in the average group it occurred between 12.5 and 13.5, and in the late group after 14. The mean ages at menarche for the three groups were, respectively, 11.7, 13.2, and 14.8.

It may be noted in Table IV and Figure 1 that early-maturing girls have a more distinct growth pattern than is true of the other two groups; their growth is extremely rapid between 11 and 13 years, and slow thereafter. The late-maturing, on the other hand, grow at a more uniform rate, with continued gains at an age when the early-maturing are gaining very little, if at all. At the age of 13 the differences between the two contrasting groups are significant at the 1 percent level,** but

*Pre-menarcheal cases are plotted with a minus age deviation from menarche; post-menarcheal cases, with a positive age deviation.

**$t = 2.75$, with 30 d.f. S.D.'s are 4.6 and 4.9 for the early and late groups respectively.

TABLE IV

MEAN SCORES (KG., RIGHT GRIP) FOR EARLY-, AVERAGE-, AND LATE-MATURING GIRLS

*Age**	*Early*	*Average*	*Late*
	N: 16	24	16
11.0	21.1	20.9	20.6
11.5	24.4	23.2	21.2
12.0	26.1	25.8	22.5
12.5	29.1	26.8	23.7
13.0	30.3	28.8	25.7
13.5	29.3	30.3	26.8
14.0	29.7	30.7	26.4
14.5	31.0	32.2	28.4
15.0	32.5	33.3	31.4
15.5	33.4	35.2	32.7
16.0	33.4	35.8	32.4
16.5	34.7	36.1	34.4
17.0	34.3	36.5	34.8
17.5	33.9	37.8	35.3

*The class interval is 10.75 to 11.24, etc.

at 14 and beyond this is no longer true.

The intense velocity of growth in the early-maturing, followed by a prompt leveling off, is paralleled by growth data on height and other physical dimensions,[8] and suggests that adolescence is a somewhat different experience for those who mature at different rates. It is probable that the physiological imbalances characteristic of a period of growth are more acute among the early- than the late-maturing, but with the latter these imbalances are, so to speak, chronic over a relatively long period. The effect upon the individual may vary widely from case to case, but in view of the interrelationship of different aspects of growth, teachers and counsellors should not neglect the possible role of precocious or delayed maturing, in connection with problems of adolescent adjustment.

After 13 the scores for the early group tend to drop below the average, although not to a statistically significant degree. Such a drop, however, is not unexpected in view of previous studies of growth in physical size. In the same sample, Bayley has found that whereas early-maturing boys are still above average in height at age 17, early-maturing girls fall lower than the average at this age.[2, p. 55] Similarly, Shuttleworth has reported that girls with the earliest ages of maximum growth

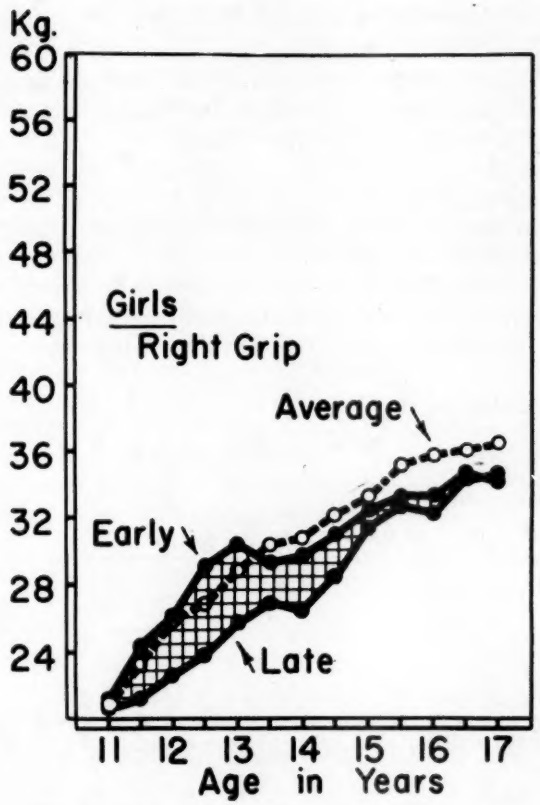

FIG. 1. Growth Curves (Right Grip) for Early-, Average-, and Late-maturing Girls, Classified on the Basis of Age at Menarche.

are the tallest in the Harvard Growth Study up to about 13 years, but are the shortest at 17.[9, p. 11] Physical precocity in girls appears, then, to be associated with an early arrest in growth; the same is not true of boys.

It is apparent that late-maturing girls lag behind the average at all ages. The difference is small, but its consistency may support the hypothesis that among girls (as among boys) the constitutional factor of ectomorphy, with relatively weak muscular development, is associated with retarded growth.[6]

In spite of the fact that growth curves in strength among the early- and late-maturing are similar to those in physical dimensions, the question may be raised as to whether the awareness of maturing, in post-menarcheal cases, may have a sharply inhibitory effect upon motor performance, perhaps by way of altered motivation. A careful study of individual records fails to reveal evidence of this, although it is probably

true that in later adolescence girls are, in general, less readily motivated than boys in physical activity tests.

An independent determination of rate of maturing is available from x-rays of the hand and knee, which were taken semi-annually and assessed by Dr. Nancy Bayley with reference to the Todd skeletal age norms.* From these cumulative records, two contrasting groups of early- and late-maturing girls were selected. As in the case of the groups chosen by the menarcheal age criterion, they do not consist of clinical extremes, but are merely arbitrary truncations of a normal distribution of school children. The early-maturing represent approximately the earliest 20 percent in a normal public school population; the late-maturing include approximately the 20 percent slowest to mature.

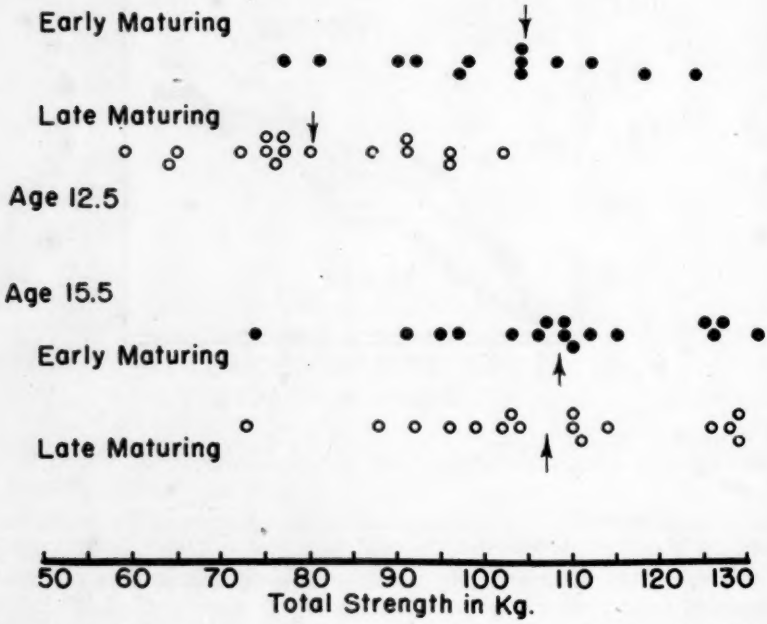

FIG. 2. Distributions, at Two Ages, of Total Strength Scores for Early- and Late-maturing Girls, Classified on the Basis of Skeletal X-Rays.

Figure 2 presents distributions of total strength (Right and Left Grip, Pull, Thrust) for 16 early- and 16 late-maturing girls (by the criterion of skeletal age). Each dot or circle represents an individual, with the mean of each distribution indicated by an arrow. In the upper half of the figure, for age 12.5 years, the difference in distributions is so marked that none of the late-maturing girls reach the mean of the early-maturing. In the lower half of the figure, at age 15.5, this differ-

*The technique of x-ray assessments, in this study, is described in reference No. 2.

ence has disappeared, and the early- and late-maturing have a closely similar spread of scores.

A comparison of Figures 1 and 2 leads to similar inferences as to the strength characteristics of different maturity groups. The contrasting samples of cases reach about the same strength level after the age of 15, but by very different routes.

An additional method of comparing age changes is through the use of standard scores. Relative strength levels, and changes in these levels, are brought out more clearly when strength measures in physical units are transformed into scale values (relative to a group mean of 50 and a group S.D. of 10). These computations have been made for the early and late groups, and the results are shown in Table V and Figure 3. Figure 3 provides graphic evidence of an increasing relative growth rate of the early-maturing girls from 11 to 13, followed by a terminal slowing-up.** The late-maturing girls exhibit a growth phase which is not shown for the early-maturing, because our data begin at too late an age to include it; this is the recession phase preceding the puberal growth spurt, a characteristic slowing in the rate of growth,

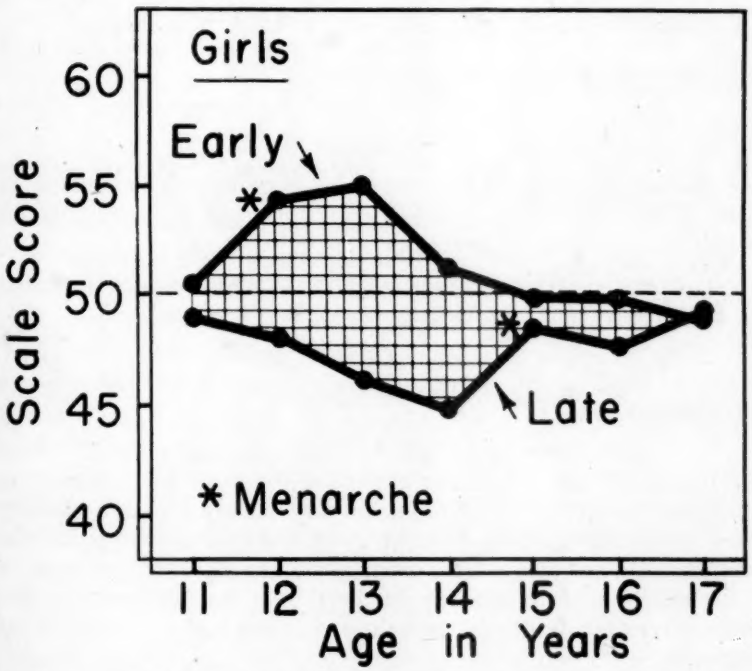

FIG. 3. Standard Score Curves (Right Grip) for Early-and Late-maturing Girls, Classified on the Basis of Age at Menarche.

**A rise in a standard score curve indicates an increasing rate of growth, relative to the total sample; a fall in the curve indicates slower-than-average growth.

which has been noted also in pre-puberal growth curves for physical dimensions.[8] The recession phase is followed by an increasing relative growth rate after 14. Among both early- and late-maturing the growth pattern has a definite relation to the menarche, which marks the approximate termination of the puberal growth spurt in strength.

TABLE V

MEAN STANDARD SCORES FOR EARLY- AND LATE-MATURING GIRLS (RIGHT GRIP)

*Age**	*Early*	*Late*
	N: 16	16
11.0	50.4	48.9
12.0	54.2	47.0
13.0	54.9	46.1
14.0	51.2	44.7
15.0	49.8	48.5
16.0	49.8	47.6
17.0	48.9	49.3

*The class interval covers half a year, 10.75 to 11.24, etc.

SUMMARY

The effect of physiological maturing upon growth in static dynamometric strength was examined by several methods in a normal sample of girls in a public school system.

1. Post-menarcheal girls are stronger than pre-menarcheal girls of the same age.

2. As shown by correlations, strength appears to be more closely related to physiological age (age deviation from the menarche) than to chronological age.

3. Growth curves in strength for groups classified as early, average, and late, in the age at reaching menarche, show that the late-maturing lag behind the average at all ages; the early-maturing have a tendency toward more rapid growth from 11 to 13, but their growth is retarded relatively soon and they reach a terminal position below the group mean. These results tend to support the hypothesis that endocrine factors which promote early puberal growth in girls also lead to an early arrest of growth.

4. Comparisons of distributions for early- and late-maturing (classified on the basis of skeletal age) show conspicuous differences at age 12.5, which is close to the time of maximum superiority of the early-maturing. With a continued slow but persistent growth of the late-

maturing, these differences disappear in the course of the next three years.

5. As shown in standard score curves, the pre-puberal growth spurt in strength begins about one year before the menarche, and reaches a peak close to the time of menarche.

6. The early- and late-maturing reach similar strength levels in later adolescence, but through different patterns of growth.

Since the physiological factors which underly these contrasting growth patterns may be expected to find expression in many other aspects of adolescent development, it is a reasonable inference that measurement and guidance in physical education should be concerned not merely with individual differences in status at a given time, but also with individual differences in the process of maturing.

BIBLIOGRAPHY

1. Bayley, N., "Skeletal X-Rays As Indicators of Maturity," *J. Consult. Psychol.*, IV, 69-73 (1940).

2. ....., "Size and Body Build of Adolescents in Relation to Rate of Skeletal Maturing," *Child Develop.*, XIV, 47-90 (1943).

3. Caldwell, M. E. *Studies On the Strength of Adolescent Girls.* Berkeley: University of California, 1940. (M.A. thesis).

4. Jones, H. E., "Principles and Methods of the Adolescent Growth Study," *J. Consult, Psychol.*, III, 157-159 (1939).

5. ....., "Procedures of the Adolescent Growth Study," *J. Consult. Psychol.*, III, 177-180, (1939).

6. ..... *Motor Performance and Growth.* Stanford University, California: Stanford University Press, 1947 (in press).

7. Reynolds, E. L., "Sexual Maturation and the Growth of Fat, Bone, and Muscle in Girls," *Child Develop.*, XVII, 121-144 (1946).

8. Shuttleworth, F. K., "Sexual Maturation and the Physical Growth of Girls Age Six to Nineteen," *Monog. Soc. Res. Child Develop.*, II:5 (1937). Pp. 252.

9. ....., "Physical and Mental Growth of Boys and Girls Ages Six Through Nineteen in Relation to Age of Maximum Growth," *Monog. Soc. Res. Child Develop.*, IV:3 (1939). Pp. vi + 291.

The History of School Health Education In West Virginia, 1863-1945

By GLADYCE H. BRADLEY
Morgan State College
Baltimore, Maryland

(Submitted for publication, January, 1947)

"WITHOUT history," says Juan Luis Vives, the Spanish scholar and writer, "no one would know anything about his father and ancestors, no one would know his own rights or those of another." It is necessary to know the history of school health education to understand and appreciate the increasing importance of child health, the efforts that have been made by various groups and individuals to protect the source of power of the nation, and the vast amount that remains to be done in the field of school health education.

This study is concerned with the development of school health education in West Virginia, beginning with 1863, the year of the state's formation and ending with 1945. By separating the early days of the state's existence, those of World Wars I and II, the time of the introduction of significant trends, developments or changes in health concepts, the time span of eighty-two years has been divided into six periods around which the data are presented. Sources of data used include the reports of the state suprintendent of schools, State Department of Education, public health laws, school laws, code of West Virginia, and additional sources which are included in the bibliography. These reports are authoritative in that they were prepared by responsible officers of the state, and complete in that they cover the entire period of educational and health activities.

Period of the Nineteenth Century, 1863-99.—When West Virginia was separated from Virginia on June 20, 1863, it was recognized that a school system was desirable. At this time there was little interest in the nation as a whole concerning health education. The increased concern about schoolhouse architecture was not based fundamentally on health improvement. Courses in physiology and hygiene, which were offered in some schools, dealt chiefly with factual information. In the vast majority of schools little attention, if any, was paid to these subjects, probably because the authorities were principally interested in providing instruction in the fundamentals of reading, writing, and arithmetic. Until these had been provided it was practically useless to think in terms of broadening the curriculum. But as time went on and the general con-

---

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Michigan, April, 1946.

ditions of the school system improved, certain factors developed which led to more interest and emphasis on school health.

The establishment of the State Board of Health on March 3, 1881,[1] perhaps did more than any other factor to develop interest in school health education in West Virginia. One of the duties of this Board was to make sanitary investigations.[2] In compliance with this regulation a complete survey of the sanitary conditions of the schools was made in 1884. Many deplorable conditions were found. Recommendations were made for the improvements of the environmental conditions of schools.[3]

Advances made in school health education during the nineteenth century were largely due to the broad vision of B. L. Butcher, state superintendent of schools, and to the general interest of the State Board of Health in more hygienic school surroundings. The initial health instruction introduced in schools was instruction on the effects of alcoholic drinks, stimulants, and narcotics on the human body. As early as 1885, State Superintendent of Schools Morgan urged that such instruction be given in all schools. This was thought to be necessary because of the widespread evil of intemperance, and the ignorance of the population about the effects of alcoholic drinks and narcotics on the body.[4] A similar opinion was shared by the State Board of Health. As a result of this interest and the activity of the Women's Christian Temperance Union, begun in 1879, the Legislature in 1887 passed an act governing instruction in alcoholic drinks and narcotics.[5] This law, like those passed by forty other states at about the same time, specified that the instruction should be a part of a broader program of instruction in physiology and hygiene. The passage of this law marked the beginning of the teaching of hygiene on a broad scale. It is interesting to note that as early as 1842 Horace Mann had pointed out the need for such instruction.

In 1878 the Legislature provided that Cutter's *Anatomy* and *Physiology* was to be used exclusively.[6] The three books to be used were Cutter's *Beginner's Anatomy, Physiology, and Hygiene,* Cutter's *Intermediate Physiology and Hygiene,* and Cutter's *Comprehensive Physiology and Hygiene.*

Instruction dealt chiefly with factual information. This is shown by the following set of questions for the third grade[7]:

Apparatus: Blackboard and chart
Subject: Our bodies
Divisions of body: Head, trunk, and limbs.

1 *Report of the State Board of Health,* 1881-2, p. 7.
2 *Ibid.,* p. 21.
3 *Biennial Report of the State Superintendent of Schools,* 1898-1900, p. 10.
4 *Ibid.,* 1885-7, p. 18.
5 *Code of the State of West Virginia,* 1887, Chap. 45, Sec. 11a.
6 *Ibid.,* 1894, Chap. 45, Art. 58.
7 *Biennial Report of the State Superintendent of Schools,* 1889-90, pp. 54-5.

The Head

1. The head is a bony box called the skull.
2. The head contains the brain.
3. The brain is the seat of thought.
4. Show how alcohol and tobacco affect the brain.

The Trunk

1. The trunk is divided into chest and stomach.
2. Explain position and office of heart and lungs.
3. Show effects of alcohol and tobacco on the heart, blood, and breath.
4. Explain the position and office of stomach.

The Limbs

1. Explain the use of bones, and joints, and muscles.
2. Show how the joints may become *stiffened* by allowing the extremities to remain cold for a length of time.
3. Show the effects of alcohol and tobacco on the limbs in unsteady gait and trembling of the hands.

Period Preceding World War I, 1900-13.—During this period school and health authorities continued to be concerned with the improvement of environmental conditions pertaining to health.[8] Likewise, there was keen interest in medical inspection of school buildings. The request by the state superintendent of schools, Morris P. Shawkey, for architectural standards for school buildings made known the desire of the state for an improved building program.[9]

Another aspect of school health education which was introduced and emphasized at this time was that of medical inspection of school children. The first function of the public schools in the field of health was protection. When laws were passed demanding that every child attend school, it became incumbent upon the state to see that the child's health was not impaired in meeting this requirement. Later it was recognized that there were conditions in the schools that offered a menace quite apart from sanitation, namely, those menaces which were chiefly in the form of communicable diseases, such as smallpox, chicken pox, measles, diphtheria, typhoid fever, and whooping cough. School authorities noticed that during the fall especially there was an epidemic of communicable diseases, directly following the opening of public schools. Communicable disease control, therefore, was the first function of early school physicians and health officers in schools.

Physicians in general, however, soon observed that some children who were not ill from communicable diseases needed medical attention. They realized that if the health of the child were to be protected it was necessary to do more than be sure that he was not ill from a communicable disease, nor made ill from the effects of an unsanitary environment.

It became necessary to establish a program commensurate with the

8 *Ibid.*, 1902-04, p. 58.
9 *Loc. cit.*

child's needs. The first part of this program was medical inspection. The legislation of 1911,[10] which required the boards of education of independent districts to provide medical inspectors and was permissive in other districts, still remains unchanged. That the law was permissive in some districts has been a great deterrent to the school health program.

Although health instruction was better organized during this period than the preceding one, the subjects and their content remained the same. However, there was a pronounced irregularity in the enrollments of courses in physiology, hygiene, and sanitation.[11] This irregularity probably was influenced by the general science movement, which began about 1900, and by the developing interest in home economics.

Period of World War I, 1913-19.—This international crisis focused the attention of the state and nation on the importance of health. Health improvement was the keynote of the period. The numerous rules and regulations which were passed were directed toward this end. In addition, the activities of the West Virginia Anti-Tuberculosis League,[12] the survey of the State Health Department by Doctor Taliaferro Clark of the United States Public Health Service,[13] and the establishment of the McDowell County Dental Clinic[14] were similarly concerned with the improvement of personal health. Furthermore, the factual material of health courses was gradually supplemented by applied health principles. Finally, physical education was recognized as a fundamental part of the school curriculum.

From the findings of medical inspection and physical examinations it became obvious that it was necessary to do something to correct the defects observed. With this extension of the health concept the second phase of the school health education program evolved, the correction of physical defects. Parents were notified of their children's defects and urged to secure the necessary corrections. School nurses were provided by boards of education and health departments to participate in these activities.

Nurses were considered particularly necessary in the follow-up program. The scope of this program has been expanded to include many important activities. It was found that some families did not have money to pay for corrections and immunizations. Therefore, it became necessary to develop free clinics. These were financed by the State Health Department, boards of education, Parent-Teacher Associations, Koppers Coal Company, and other agencies and organizations.

Period Following World War I, 1920-29.—The modern program of

---

10 Acts of the Legislature, 1911, Chap. 31, *Code of the State of West Virginia*, 1919, Chap 2, Art. 64.

11 *Report of State Supervisor of High Schools*, 1915-16, p. 17.

12 S. Adolphus Knopf. *A History of the National Tuberculosis Association*. New York: National Tuberculosis Association, 1922, p. 22.

13 *Public Health Reports*, XXX, p. 238.

14 *Biennial Report of the State Superintendent of Schools*, 1919-20, p. 81.

school health education has developed since 1917. Much of the credit for its initiation, promotion, and development in the United States and West Virginia should be given to the Child Health Organization of America (1917-29). Other private agencies, including the American Child Health Association, the National Tuberculosis Association, and the National Education Association together with governmental agencies, colleges, and universities, provided health leadership. This was evidenced in printed materials, the Health Crusade, and teacher education in health.

Health was stressed as an essential for education by the Report of the Commission on the Reorganization of Secondary Education, which in 1918 outlined the Seven Cardinal Principles. Health was listed as the first principle. Recognition was made also of the relationship of health to some of the other principles. Thus, the idea that the school's only responsibility was to teach facts was supplanted by the idea that the school's responsibility was to train for citizenship, with due consideration for the child's physical, mental, social, and emotional health and welfare.

Until the period of World War I, physical education in elementary schools had been limited to calisthenics and exercise. Many secondary schools did not offer physical education courses; others had mass athletic programs, football and basketball. During the period, 1920-29, as a result of the health deficiencies found in the draft examinations for World War I, a demand for health and physical education appeared. A statewide program was introduced under the supervision of Melville K. Stewart, who later was succeeded by Arthur G. Upton. The program was greatly aided by the publication of state manuals on health and physical education.

The outstanding difficulty of the new movement in health education in the twenties was an almost entire lack of anything approaching acceptable standards of procedure. Various specialists, in addition to regular school personnel, were working toward their own objectives, as if they had such objectives entirely independent of other workers in the field. Gradually, however, the program became better organized and coordinated as efforts were made to plan the program in terms of the needs of the individual child. This was not true of health education alone but of the entire school curriculum.

Period from 1930-39.—Increased use was made of successful experiences of the preceding period, such as medical inspection of school buildings, medical inspection of school children, interest of schools in hygiene instruction, and physical education. The effects of the child-centered idea in education were marked. Various sociological developments in relation to environment, housing, social reform, and playgrounds affected school health education. Of greatest significance was the interest of the federal government in health education which re-

sulted in the White House Conference on Child Health and Protection and crystallized interest in child health.

The school survey of 1928 pointed out the numerous school building needs and paved the way for the improvement of environmental conditions pertaining to the health of the school child. However, the establishment of rules and regulations governing standardization of schools led to better playgrounds, improved water supply, and sanitary rural school toilets.

There was teacher participation in medical inspection accompanied by an increase in the number of health supervisors. This resulted in better medical examinations. Boards of education were given authority to require all teachers to have certificates of good health.[15] This recognition of the relationship of the health of the teacher to the total school program was of great importance.

The most pronounced development, however, was shown in the preparation and publication of courses of study in health and physical education, and the effects of alcoholic drinks and narcotics on the human body. On the other hand, the abolition of the office of state director of physical education in 1932, due to lack of funds, became a handicap to school health education. This action was due to lack of funds characteristic of the depression. That such action was resented by the people is evidenced by the fact that its legality was questioned and the matter was taken to the Supreme Court. The latter ruled that the abolition of this office was unconstitutional.[16] Despite this fact the office has not yet been restored.

During the summer of 1931, representatives of the President's Organization on Unemployment Relief, following a survey of the needs in West Virginia, approached the American Friends Service Committee of Philadelphia, and requested them to undertake relief work among the children of unemployed miners in the bituminous coal fields. After very careful consideration the committee decided to undertake this service.[17] In September, 1931, feeding was begun by this committee. The peak load was reached in April, 1932, when 33,227 school children were furnished one substantial meal a day, and in addition milk was furnished to 7,697 pre-school children. This project was further developed by the Works Project Administration.

Recent Developments, 1940-45.—West Virginia schools have more recently given increased attention to health protection. Three fundamental activities in this area are hygienic arrangement of the school program to provide healthful living, the physical education program,

15 *Supplement to the Code of the State of West Virginia*, 1933, Chap. 18, Art. 1752.
16 *Biennial Report of the State Superintendent of Schools*, 1930-32, p. 18.
17 *Report of the Child Relief Work in the Bituminous Coal Fields*, 1932, p. 3.

and health instruction. Hygienic arrangement of the school program includes administrative planning of recesses, length of recitation periods, sequence of subjects, kinds of examinations, amount of homework allowed, number of children per class or per teacher, maintenance of social and wholesome atmosphere in the classroom, arrangement for lunches, and other activities in the school program that are related to the child's health.

Physical education programs have been expanded and emphasized. The High School Victory Corps Program promoted by the United States Office of Education and the redefining of the contribution of physical education to general education and to physical fitness gave impetus to this movement.

More time is being devoted to health education classes. Often these classes or their materials are integrated or correlated with physical education, science, and home economics classes. Health in schools in modern times is not a subject but a manner of living, and the experiences provided in the more recent teacher's manual lend themselves to this concept.

There was a need for special health classes for children who could not successfully carry forward the regular school program because of poor physical conditions or some physical defect. A school for the deaf, dumb, and blind was established at Romney in 1870,[18] and at Institute in 1926.[19] No legal provision was made for other handicapped children until 1941.[20] In that year funds were provided for an educational program for these children. Included in this program were classes for crippled children, speech improvement, and the mentally retarded. Funds have not been sufficient to include all children needing this kind of instruction.

The modern school health education program recognizes significant contributions of the various divisions of the State Department of Education, State Health Department, county health department, state physical education association, state nutrition committee, Department of Public Assistance, and more recently the W. K. Kellogg Foundation. These agencies and organizations in cooperation with the schools are responsible for school health education as it exists in West Virginia today.

---

18 *Code of the State of West Virginia*, 1870, Art. 8, Sec. 1855.

19 *Biennial Report of the State Superintendent of Schools*, 1924-26, p. 145.

20 *Ibid.*, 1942-4, p. 71.

BIBLIOGRAPHY

Biennial Report of the State Superintendent of Schools of West Virginia. Charleston: State Department of Education, 1863-1945.

Biennial Report of the State Board of Education of West Virginia. Charleston: State Department of Education, 1911-45.

Report of the State Board of Health of West Virginia. Charleston: State Health Department, 1881-1916.

Biennial Report of the State Department of Health of West Virginia. Charleston: State Health Department, 1919-45.

Code of the State of West Virginia.

The Public Health Laws of West Virginia. Charleston: State Health Department, 1915, 1930, 1939.

Report of the Child Relief Work in the Bituminous Coal Fields. American Friends Service Committee. Philadelphia: American Friends, 1932.

Report of State Supervisor of High Schools. Charleston: State Department of Education, 1915-16.

Knopf, S. Adolphus. *A History of the National Tuberculosis Association.* New York: National Tuberculosis Association, 1922.

Public Health Reports, XXX. Charleston: State Health Department.

Supplement to the *Code of the State of West Virginia,* 1933.

Components of Variation and the Consistency of Repeated Measurements

By Josef Brozek and Howard Alexander
Laboratory of Physiological Hygiene
University of Minnesota
Minneapolis, Minnesota

(Submitted for publication, January, 1947)

THE statistical technique referred to as "analysis of variance" has been used in biological work primarily for testing *significance of the differences between the means* of two or more sets of measurements by the *F* ratio. In addition, in a set of measurements repeated on a group of individuals, the technique can be used for estimating the magnitude of the *components of variation* and provides means for determining the *consistency* with which a variable is measured in a series of repeated testing trials.

The procedures and assumptions involved in the statistical analysis of the sources of variation have been discussed by Snedecor[4*] and Crump.[3] The problem of reliability has been treated in detail by Alexander.[1] In this paper we shall present the fundamental concepts and the rationale of the methods, develop the formulae, and give the computational procedures. An attempt will be made to treat the subject in a way understandable to a scientist who does not have a technical training in statistics.

1. *The concept of variance.*—In describing a series of measurements one would like to know at least two general characteristics of the series: the mean value and the variability of the measurements. Variability is frequently expressed in terms of variance. The variance of a series containing n measurements is equivalent to the standard deviation squared and is defined by:

$$V = \frac{\text{sum of squared deviations from the mean}}{\text{degrees of freedom}} \qquad (1)$$

The "degrees of freedom" in this case represent the number of measurements minus one.

The degrees of freedom represent the minimum number of measurements required in order to reconstruct the series

Mr. Alexander is at present at Adrian College, Adrian, Michigan.

* Superior figures refer to numbered bibliography at end of article. In reference No. 4 note particularly pages 228-261.

when the mean is given. If the mean of three measurements equals 5, then, knowing that two of the measurements are 4 and 3, the third one can easily be calculated and is found to be 8. Thus in a series of n measurements the degrees of freedom are (n—1).

Defining y, a deviation of a measurement (Y) from the mean of the series ($\overline{Y}$), as $y=Y-\overline{Y}$, we can write

$$V=\frac{\sum_{1}^{n} y^2}{n-1} \qquad (1a)$$

$$V=\frac{\sum_{1}^{n} (Y-\overline{Y})^2}{n-1} \qquad (1b)$$

2. *Score components in a two-way table.*—Two sets of means can be calculated for the data from experiments in which n individuals have been measured on k days: (1) The means for the *individuals,* each individual being considered over the period of k days, and (2) the means for the *days,* these daily means being based on n individuals. The arrangement is represented in Table I.

TABLE I

SYMBOLIC REPRESENTATION OF A TWO-WAY TABLE WITH k COLUMNS (DAYS) AND n ROWS (INDIVIDUALS)

		Days						Individual Totals $\sum_{1}^{k} Y_I$	Individual Totals $\overline{Y}_I$
		1	2	—	—	—	k		
Individuals	1	Y_{11}	Y_{12}	—	—	—	Y_{1k}	T_{I1}	$\overline{Y}_{I1}$
	2	Y_{21}	Y_{22}	—	—	—	Y_{2k}	T_{I2}	$\overline{Y}_{I2}$
	|	|	|				|	|	|
	|	|	|				|	|	|
	|	|	|				|	|	|
	n	Y_{n1}	Y_{n2}	—	—	—	Y_{nk}	T_{In}	$\overline{Y}_{In}$
Daily Totals $\sum_{1}^{n} Y_D$		T_{D1}	T_{D2}	—	—	—	T_{Dk}	T	
Daily Means $\overline{Y}_D$		$\overline{Y}_{D1}$	$\overline{Y}_{D2}$	—	—	—	$\overline{Y}_{Dk}$		$\overline{Y}$

Y=measurement on any individual on any day.
Y_{11}=measurement on the first individual on the first day, etc.
T_{I1}=sum of the k daily measurements on the first individual.

T_{D1}=sum of n individual measurements on the first day.
T=sum of all values on all days.
$\overline{Y}$=mean of all values on all days.

In a two-way table a single measurement or a "score," Y, obtained for a given individual on a given day, can be regarded as being composed of four parts:

$$Y=\overline{Y}+(\overline{Y}_I-\overline{Y})+(\overline{Y}_D-\overline{Y})+[Y-\overline{Y}-(\overline{Y}_I-\overline{Y})-(\overline{Y}_D-\overline{Y})] \qquad (2)$$

$$Y=\overline{Y}+\ i\ +\ d\ +\ r \qquad (2a)$$

where Y=a measurement

$\overline{Y}$=mean of all n.k measurements, $\frac{\sum_{1}^{nk} Y}{n.k}$

$\overline{Y}_I$=mean of k measurements on an individual, $\frac{\sum_{1}^{k} Y_D}{k}$

$\overline{Y}_D$=mean of n measurements obtained on a given day, $\frac{\sum_{1}^{n} Y_D}{n}$

The four terms on the right side of the equation (*2a*) represent the following components of a measurement:

$\overline{Y}$ is a component *common* to all n.k. measurements and indicates the overall level of the scores. Mathematically it is defined as the general mean;

i characterizes the average level of the measurements obtained on an *individual* and is defined as the deviation of the individual mean from the general mean, $i=\overline{Y}_I-\overline{Y}$;

d characterizes the average level of the measurements obtained on a given *day* and is defined as the deviation of the day's mean from the general mean, $d=\overline{Y}_D-\overline{Y}$; and

r is the residual or *random* component which remains after we have accounted for the portions of the score contributed by the general mean, the individual mean, and the daily mean,

$$r=[Y-\overline{Y}-(\overline{Y}_I-\overline{Y})-(\overline{Y}_D-\overline{Y})]$$

From equation *2a* the following transposition can be made:

$$Y-\overline{Y}=i+d+r \qquad (2b)$$

$$y=i+d+r \qquad (2c)$$

which indicates that the deviation of a score from the general mean ($y=Y-\overline{Y}$) is a sum of the components i, d, and r.

The scores in Table II, simulating a set of measurements on five individuals tested on four days, are analyzed into their components as indicated in Table III. Each score Y can be obtained as the sum of the components $\overline{Y}$, i, d, and r. The differences, y, between the measurements Y (Table I) and the general mean $\overline{Y}$ (Table III 1) are given in Table IV.

TABLE II

SET OF FICTITIOUS MEASUREMENTS, Y, ON n INDIVIDUALS OVER k DAYS

		Days 1	2	3	4	T_I	$\overline{Y}_I$
Individuals	1	30	32	28	30	120	30
	2	19	32	35	30	116	29
	3	20	38	23	23	104	26
	4	33	36	29	38	136	34
	5	28	27	40	29	124	31
T_D		130	165	155	150	600	
$\overline{Y}_D$		26	33	31	30		30

TABLE III

COMPONENTS OF THE MEASUREMENTS Y FROM TABLE II

$Y=\overline{Y}+i+d+r$

Individuals	1. Component $\overline{Y}$ Days 1	2	3	4	2. Component i Days 1	2	3	4	3. Component d Days 1	2	3	4	4. Component r Days 1	2	3	4
1	30	30	30	30	0	0	0	0	−4	3	1	0	4	−1	−3	0
2	30	30	30	30	−1	−1	−1	−1	−4	3	1	0	−6	0	5	1
3	30	30	30	30	−4	−4	−4	−4	−4	3	1	0	−2	9	−4	−3
4	30	30	30	30	4	4	4	4	−4	3	1	0	3	−1	−6	4
5	30	30	30	30	1	1	1	1	−4	3	1	0	1	−7	8	−2

TABLE IV

DEVIATIONS OF SCORES IN TABLE II FROM THE GENERAL MEAN

$y=Y-\overline{Y}=(i+d+r)$

Individuals	Days 1	2	3	4
1	0	+2	−2	0
2	−11	+2	+5	0
3	−10	+8	−7	−7
4	+3	+6	−1	+8
5	−2	−3	+10	−1

The sum of the squares of the n.k values of y in Table IV can be obtained by adding up the sums of i^2, d^2, and r^2:

$$\sum_1^{nk} y^2=\sum_1^{nk} i^2+\sum_1^{nk} d^2+\sum_1^{nk} r^2 \qquad (3)$$

$$644 = 136+ 130+378$$

The "degrees of freedom" (nk—1), for the n.k values of y also can be partitioned into three quantities corresponding, respectively, to i, d, and r:

$$nk-1=(n-1)+(k-1)+(n-1)\,(k-1) \qquad (4)$$

3. *Variances in a two-way table.*—On the basis of the values i, d, r, and y we can obtain in a two-way table four variances, sometimes referred to as "mean squares":

(1) A variance based on the differences *between individuals:*

$$V_{bI}=\frac{\sum_{1}^{nk} i^2}{n-1} \qquad (5)$$

(2) A variance based on the differences *between days:*

$$V_{bD}=\frac{\sum_{1}^{nk} d^2}{k-1} \qquad (6)$$

(3) A variance based on the *random* components of the scores:

$$V_{R}=\frac{\sum_{1}^{nk} r^2}{(n-1)(k-1)} \qquad (7)$$

(4) A variance representing variation of the scores, irrespective of their classification according to individuals or days, or a *total* variance in the table:

$$V_{T}=\frac{\sum_{1}^{nk} y^2}{nk-1} \qquad (8)$$

$$V_{T}=\frac{\sum_{1}^{nk} i^2+\sum_{1}^{nk} d^2+\sum_{1}^{nk} r^2}{(n-1)+(k-1)+(n-1)(k-1)} \qquad (8a)$$

The variance V_T represents a derived term, and is of little interest as compared with the first three variances, which were obtained independently from each other.

4. *Simple computational formulae for variances in a two-way table.*—The scheme indicated in Table V greatly simplifies computation of the terms V_{bI}, V_{bD}, and V_R. This scheme is to be used in conjunction with the data as set up in Tables II and III.

TABLE V

SCHEME FOR OBTAINING VALUES NEEDED FOR COMPUTATION OF THE VARIANCES V_{bI}, V_{bD}, AND V_R

$\Sigma_1=\sum_{1}^{nk} Y^2$ $N_1=nk$	$\Sigma_2=\frac{\sum_{1}^{n}(T_I)^2}{k}$ $N_2=n$
$\Sigma_3=\frac{\sum_{1}^{k}(T_D)^2}{n}$ $N_3=k$	$\Sigma_4=\frac{(T)^2}{nk}$ $N_4=1$

where Y=any score

$T_I=\sum_{1}^{k} Y_I$=sum of the k scores for an individual

$T_D=\sum_{1}^{n} Y_D$=sum of the n scores for a day

$T=\sum_{1}^{nk} Y$=grand total

The numerator in the variance formulae is obtained by adding or subtracting the sums Σ_1, Σ_2, Σ_3, and Σ_4, drawn from Table V; the denominator, by following the same pattern of additions or subtractions of the corresponding numbers, N_1, N_2, N_3, and N_4. The operations involved are as follows:

$$V_{bI}=\frac{\Sigma_2-\Sigma_4}{N_2-N_4} \qquad (9)$$

$$V_{bD}=\frac{\Sigma_3-\Sigma_4}{N_3-N_4} \qquad (10)$$

$$V_R=\frac{\Sigma_1-\Sigma_2-\Sigma_3+\Sigma_4}{N_1-N_2-N_3+N_4} \qquad (11)$$

By substituting the terms represented by the symbols Σ and N in the above equations, it can be shown that they are identical with the analytical formulae (5), (6), (7), and (8).

The data in Table II, obtained from random two-digit numbers between 19 and 40, yield values of Σ and N as given in Table VI.

TABLE VI

VALUES OF Σ AND N, AND THE VARIANCES V_{bI}, V_{bD}, AND V_R FOUND FOR THE DATA FROM TABLE II

$\Sigma_1=18{,}644$ $N_1=5\times4=20$	$\Sigma_2=\frac{72{,}544}{4}=18{,}136$ $N_2=5$
$\Sigma_3=\frac{90{,}650}{5}=18{,}130$ $N_3=4$	$\Sigma_4=\frac{(600)^2}{20}=18{,}000$ $N_4=1$

$$V_{bI}=\frac{18{,}136-18{,}000}{5-1}=\frac{136}{4}=34.0$$

$$V_{bD}=\frac{18{,}130-18{,}000}{4-1}=\frac{130}{3}=43.3$$

$$V_R=\frac{18{,}644-18{,}136-18{,}130+18{,}000}{20-5-4+1}=\frac{378}{12}=31.5$$

Using the F ratios, we can test whether the variances between individuals and days (V_{bI} and V_{bD}) are significantly larger than the random variance:

$$F_{bI}=V_{bI}/V_R \qquad (12)$$

$$F_{bD}=V_{bD}/V_R \qquad (13)$$

Substituting the values from Table VI we have

$$F_{bI}=34.0/31.5=1.079$$
$$F_{bD}=43.3/31.5=1.374$$

The variances entering into the F tests are associated with the following degrees of freedom (d.f.): for V_{bI}, d.f.=n-1=4; for V_{bD}, d.f.=k-1=3; for V_R, d.f.=(n-1) (k-1)=12. Snedecor's Table 10.7[4, p. 222] gives the 5 percent level of significance for an F associated with 4 and 12 d.f. as 3.26; for 3 and 12 d.f., the 5 percent point of F is 3.49. Therefore, neither of the two F values obtained above reach the level of statistical significance. This means that in Table II there is only one important source of variation, the random fluctuation of the scores.

5. *Composite total variance.*—The terms V_{bI}, V_{bD}, and V_R, derived from a given sample of n individuals tested on k days, may be used for obtaining an estimated total variance in a population of an infinite number of such individuals measured over an infinite number of days. This composite estimate of "population variance" can be expressed in terms of its components as follows:

$$\sigma^2=\sigma_I^2+\sigma_D^2+\sigma_R^2 \qquad (14)$$

where σ^2=composite estimate of the total variance,
σ_I^2=component due to the differences between the test values obtained by different individuals
σ_D^2=component resulting from differences between test values obtained on different days (trials), and
σ_R^2=component representing the random fluctuation of the repeated measurements.

The components cannot be determined directly and must be estimated on the basis of the given sample or samples studied. The estimate of one of the above terms can be obtained as :

$$\sigma_R^2=V_R \qquad (15)$$

V_R is computed as indicated in formulae 7 and 11.

The values σ_I^2 and σ_D^2 are obtained from the terms V_{bI} and V_{bD}, to which they are related as follows:

$$V_{bI}=k\cdot\sigma_I^2+\sigma_R^2 \qquad (16) \qquad\qquad V_{bD}=n\cdot\sigma_D^2+\sigma_R^2 \qquad (17)$$

where k=the number of days and n=the number of individuals.

These equations lead to:

$$\sigma_I^2=\frac{V_{bI}-V_R}{k} \qquad (18) \qquad\qquad \sigma_D^2=\frac{V_{bD}-V_R}{n} \qquad (19)$$

The values σ_I^2, σ_D^2, and σ_R^2 indicate the estimated magnitude of the individual, daily, and random component of the total variance.

For the data of Table II we obtain the following values:

$$\sigma_R^2 = V_R = 31.5$$

$$\sigma_I^2 = \frac{V_{bI} - V_R}{k} = \frac{34.0 - 31.5}{4} = 0.6$$

$$\sigma_D^2 = \frac{V_{bD} - V_R}{n} = \frac{43.3 - 31.5}{5} = 2.4$$

The composite estimate of total variance σ^2, according to formula (*14*), is

$$\sigma^2 = 0.6 + 2.4 + 31.5 = 34.5$$

Random component contributes the largest part, 91.3 percent; the individual component, 1.7 percent; and the daily component, 7.0 percent of the total.

6. *Average intra-individual and inter-individual variance.* In characterizing a given set of n individuals studied over a period of k days, we can use advantageously an average intra-individual variance (i.e., variance between daily measurements *within individuals,* V_{wI}) and an average inter-individual variance (i.e., variance between individual measurements *within days,* V_{wD})

$$V_{wI} = \frac{\sum_1^n \frac{\sum_1^k (Y_I - \bar{Y}_I)^2}{k-1}}{n} \qquad (20)$$

$$V_{wD} = \frac{\sum_1^k \frac{\sum_1^n (Y_D - \bar{Y}_D)^2}{n-1}}{k} \qquad (21)$$

The term V_{wI} indicates the average magnitude of day-to-day variation in the scores; V_{wD} represents the average amount of individual - to - individual variation for the given sample of subjects.

Using symbols defined in Section 4, the simple computational formula for V_{wI} and V_{wD} are:

$$V_{wI} = \frac{\Sigma_1 - \Sigma_2}{N_1 - N_2} \qquad (22)$$

$$V_{wD} = \frac{\Sigma_1 - \Sigma_3}{N_1 - N_3} \qquad (23)$$

Using the values given in Table VI:

$$V_{wI} = \frac{18{,}644 - 18{,}136}{20 - 5} = \frac{508}{15} = 33.9$$

$$V_{wD} = \frac{18{,}664 - 18{,}130}{20 - 4} = \frac{514}{16} = 32.1$$

The variances V_{wI} and V_{wD} can be thought of as the sum of two components:

$$V_{wI}=\sigma_D^2+\sigma_R^2 \qquad (24)$$
$$V_{wD}=\sigma_I^2+\sigma_R^2 \qquad (25)$$

Because of their simplicity, these equations may be used with profit for obtaining the values σ_D^2 and σ_I^2, two of the estimated components of the composite total variance:

$$\sigma_D^2=V_{wI}-\sigma_R^2=V_{wI}-V_R \qquad (26)$$
$$\sigma_I^2=V_{wD}-\sigma_R^2=V_{wD}-V_R \qquad (27)$$

Equations (*26*) and (*27*) facilitate a direct, intuitive understanding of the terms σ_D^2 and σ_I^2: they represent the amount of variation arising from the differences between daily scores and individual scores, respectively, after the random variation in the scores has been removed.

Using the figures for a given sample we obtain:

$$\sigma_D^2=33.9-31.5=2.4$$
$$\sigma_I^2=32.1-31.5=0.6$$

The values check with those obtained by formulae (*18*) and (*19*).

7. *Criteria of consistency of repeated measurements based on random variance.*—The random variance, V_R, is a measure of inconsistency in the measurements repeated on separate days (or trials). The random component, r, of each score is the discrepancy between the actual score, Y, and the predicted score, $\widehat{Y}$

$$r=Y-\widehat{Y} \qquad (28)$$

where $\widehat{Y}$ is obtained as a sum of the grand mean, $\overline{Y}$, the deviation of an individual mean ($\overline{Y}_I$) from the grand mean, and the deviation of a day's mean ($\overline{Y}_D$) from the grand mean:

$$\widehat{Y}=\overline{Y}+(\overline{Y}_I-\overline{Y})+(\overline{Y}_D-\overline{Y}) \qquad (29)$$

The larger the discrepancies between the expected ($\widehat{Y}$) and the actual (Y) test scores, the greater the fluctuations in the individual's standing with respect to the performance of the group. This lack of consistency is reflected in the magnitude of the random variance, V_R. As indicated in equation (*7*), V_R is obtained as a sum of the r values squared, $\sum_1^{nk} r^2$, divided by the degrees of freedom, (N-1) (k-1).

The random variance, V_R (and its square root, $S.D._R$, the standard deviation of the random components of the test scores or the "standard random error" of the scores) is expressed in terms of the absolute units of measurement used in a particular test, such as kilograms of weight lifted, time of a run in seconds, number of taps in a tapping test, etc. The values V_R or $S.D._R$, obtained for different tests, are not directly compar-

able. Under certain conditions we may compare the consistency of performance in different tests by expressing the $S.D._R$ as the percentage of the grand mean.

It is much more satisfactory to evaluate the magnitude of the random variation in reference to the variability of the scores obtained by the different individuals. The best estimate of the amount of inter-individual variation is the value V_{wD}, the average inter-individual variance or the variance within days, defined in the equation (*21*).

The comparison may take a form of a simple percentage ratio:

$$V_R\% = \frac{V_R}{V_{wD}} \times 100 \qquad (30)$$

or of a coefficient of consistency, r_C:

$$r_C = 1 - \frac{V_R}{V_{wD}} \qquad (31)$$

This can also be written as

$$r_C = \frac{V_{wD} - V_R}{V_{wD}} \qquad (31a)$$

When $V_R = O$, i.e., when there is no random variation in the scores, the value of the coefficient of consistency $r_C = 1$; this means that the scores obtained by different individuals on different days are identical or differ from day to day by a factor which is the same for all individuals. The larger the V_R, at the same V_{wD}, the closer the r_C approaches zero. When $V_R = V_{wD}$, $r_C = 0$.

This consistency coefficient is determined, as indicated in formula (*30*), by two factors: (1) the amount of random variation, measured as V_R, and (2) the amount of inter-individual variation, expressed as V_{wD}. It may be helpful to repeat that V_{wD} is the variance between the scores obtained by the individuals comprising the sample, averaged for the k days on which the measurements were repeated. In laboratory investigations of fitness, the subjects usually represent a very homogeneous group with a narrow spread in age, health, performance capacity, etc. The greater the homogeneity, i.e., the smaller the differences between individuals comprising the sample studied, the lower—for the same magnitude of random variation in the scores—will be the coefficient of consistency, r_C.

To evaluate correctly the magnitude of the random error we should know the "true" variance of the individual scores in the population. No amount of statistical gymnastics can provide this datum. The information can be obtained only by

testing a number of individuals large enough to yield a truly representative sample of the population. When we are concerned with tests in which training does not affect the inter-individual variance in successive trials, the task is well within the reach of realization. Where the inter-individual variance does change with practice, the obstacles to obtaining an inter-individual variance for the population are discouraging.

The consistency coefficient, r_C, can be expressed in terms of the estimated components of the total composite variance:

$$r_C = \frac{\sigma_I^2}{\sigma_I^2 + \sigma_R^2} \qquad (32)$$

It is apparent that the daily component, σ_D^2, of the total composite variance does not enter in; the value of r_C is not affected by the variations in the daily (trial) means and can, therefore, be used for evaluation of the consistency of repeated measurements in which a trend is present.

The r_C yields for a series of repeated trials a consistency coefficient comparable to the ordinary product-moment coefficient of correlation, used when only two test trials are available. If we assume that the inter-individual variances in successive testing trials are homogeneous, then the r_C represents the average of the product-moment correlation coefficients computed for the k testing trials, taken two at a time; there are $\frac{k(k-1)}{2}$ such pairs, yielding an equal number of coefficients of product-moment correlation.

The formula for r_C can be written in a form which brings out the relationship of r_C to the coefficient of intra-class correlation[5, p. 245]

$$r_C = \frac{V_{bI} - V_R}{V_{bI} + (k-1)\,V_R} \qquad (33)$$

where k=number of trials (days) and V_{bI}=variance "between individuals," defined in equation (5).

Using the values obtained in our fictitious sample of 5 individuals measured on 4 days we obtain:

$$r_C = 1 - \frac{V_R}{V_{wD}} = 1 - \frac{31.5}{32.1} = 1 - 0.981 = 0.019$$

$$r_C = \frac{\sigma_I^2}{\sigma_I^2 + \sigma_R^2} = \frac{0.6}{0.6 + 31.5} = \frac{0.6}{32.1} = 0.019$$

$$r_C = \frac{V_{bI} - V_R}{V_{bI} + (k-1)\,V_R} = \frac{34.0 - 31.5}{34.0 + (4-1)\,31.5} = \frac{2.5}{131.0} = 0.019$$

8. *Statistical significance of a coefficient of consistency,* r_C.—The statistical significance of a given coefficient of consistency is obtained in the form of an F test:

$$F_{r_C}=\frac{V_{bI}}{V_R} \qquad (34)$$

The values for data in Table II yield

$$F_{r_C}=\frac{34.0}{31.5}=1.079$$

associated with 4 and 12 degrees of freedom. The corresponding 5 per cent point for the F test is 3.26. Our F_{r_C} is far smaller and cannot be considered statistically significant.

9. *Statistical significance of a difference between two values of* r_C.—In comparing the consistency of two different functions, and of the same function measured under different conditions or with different techniques, we want to know whether the difference between the two values of r_C is statistically significant. The procedure for making such comparisons will be essentially identical with that provided by Fisher.[4, 234ff]

The steps are (1) to determine the range in which the values of r_C could be expected to fall due simply to random sampling, and (2) to examine whether or not the ranges corresponding to the given two values of r_C overlap. Overlap of the ranges would indicate that the two values could be thought of as characteristics of two samples drawn from the *same* population, i.e., the difference would not be considered as statistically significant.

10. *Practical significance of a given* r_C.—No definite limits have been established for evaluating the practical significance of the consistency coefficient. The consistency demanded from a test will depend in part on the purpose for which the data will be used, and useful generalizations will emerge only from the actual study of the consistency of various functions measured in connection with experimental studies on human fitness.

Provided that the group of subjects under study is reasonably heterogeneous with respect to its fitness characteristics, tentative evaluative limits may be set as follows: $0.91 \leqq r_C \leqq 1.00$ would indicate a very high consistency of the repeated measurements, $0.81 \leqq r_C \leqq 0.90$ high consistency, $0.71 \leqq r_C \leqq 0.80$ moderately high consistency, and $r_C \leqq 0.70$ low consistency.

11. *Examples.*—In order to make the presentation more concrete, two sets of actual data will be analyzed in terms of the methods described in the preceding sections. The data included in Table VII represent test scores obtained by the

individual subjects in a two-plate tapping test and in a test of coordination. The details of testing methods and the experimental regimen were given elsewhere.[2]

TABLE VII

SET OF ACTUAL DATA OBTAINED ON EIGHT SUBJECTS MEASURED IN REPEATED TESTING SESSIONS

A. Tapping Speed, total number of taps in 10 seconds

Subjects	*Days of the Experiment* 45	78	117	141	T_I	$\sum_{1}^{4} Y_I^2$
G	62	64	65	65	256	16,390
W_1	74	72	69	72	287	20,605
W_2	79	80	75	75	309	23,891
T	75	75	67	70	287	20,639
N	72	76	72	75	295	21,769
S	73	74	71	71	289	20,887
Jo	67	72	72	75	286	20,482
Ja	67	77	69	66	279	19,535
T_D	569	590	560	569	2288	
$\sum_{1}^{8} (Y_D)^2$	40,677	43,670	39,270	40,581		164,198

$\sum_{1}^{8} (T_I)^2 = 655{,}938$

$\sum_{1}^{4} (T_D)^2 = 1{,}309{,}222$

$T^2 = 5{,}234{,}944$

B. Coordination, number of errors

Subjects	*Days of the Experiment* 45	78	117	141	T_I	$\sum_{1}^{4} Y_I^2$
G	33	33	22	17	105	2,951
W_1	26	29	27	21	103	2,687
W_2	16	16	12	11	55	777
T	35	25	27	25	112	3,204
N	35	29	30	20	114	3,366
S	37	31	26	16	110	3,262
Jo	43	30	25	24	122	3,950
Ja	52	40	43	26	161	6,829
T_D	277	233	212	160	882	
$\sum_{1}^{8} (Y_D)^2$	10,393	7,113	6,136	3,384		27,026

$\sum_{1}^{8} (T_D)^2 = 201{,}562$

$\sum_{1}^{4} (T_I)^2 = 103{,}104$

$T^2 = 777{,}924$

Table VII contains both the raw data and the derived totals. The Σ and N values needed for the computation of the sample variances and the derived values are given in Table VIII.

TABLE VIII

COMPUTATION OF SAMPLE VARIANCES, ESTIMATED VARIANCE COMPONENTS, F-TESTS OF SIGNIFICANCE OF THE DIFFERENCES BETWEEN INDIVIDUALS AND BETWEEN DAYS, AND COEFFICIENTS OF CONSISTENCY

A		B	
Tapping Speed		*Coordination*	
Σ_1=164,198	Σ_2=163,984	Σ_1=27,026	Σ_2=25,776
N_1=32	N_2=8	N_1=32	N_2=8
Σ_3=163,653	Σ_4=163,592	Σ_3=25,195	Σ_4=24,310
N_3=4	N_4=1	N_3=4	N_4=1
V_{bI}=392/7 = 56.00		V_{bI}=1,466/7 =209.43	
V_{bD}= 61/3 = 20.33		V_{bD}= 885/3 =295.00	
V_R=153/21 = 7.29		V_R= 365/21 = 17.38	
V_{wI}=214/24 = 8.92		V_{wI}=1,250/24 = 52.08	
V_{wD}=545/28 = 19.46		V_{wD}=1,831/28 = 65.39	
σ_I^2= 12.17	57.7%	σ_I^2= 48.01	48.0%
σ_D^2= 1.63	7.7%	σ_D^2= 34.70	34.7%
σ_R^2= 7.29	34.6%	σ_R^2= 17.38	17.3%
σ^2= 21.09	100 %	σ^2=100.09	100 %
F_{bI}= 56.00/7.29= 7.68		F_{bI}=209.43/17.38= 12.05	
F_{bD}= 20.33/7.29= 2.78		F_{bD}=295.00/17.38= 16.97	
r_C= 0.625		r_C= 0.734	

In the tapping test we have only one source of variation which is significantly greater than the random error, namely the differences between individuals. The coefficient of correlation appears to be low, even though it is statistically significant.

The low value is due primarily to the extreme homogeneity of the individuals comprising the sample. The standard random error of the repeated measurements is small ($S.D._R=\sqrt{V_R}=2.7$), representing only 3.8 percent of the mean score ($\overline{Y}=71.5$).

The scores in the cordination test exhibit a definite trend toward lower values as the time goes on; both "days" and "individuals" are significant sources of variation. The coefficient of correlation is somewhat higher than in the tapping test.

12. *Criteria of consistency of repeated measurements based on the variance within individuals.*—The variance within individuals, V_{wI}, and its square root, $S.D._{wI}$, the average standard

deviation of daily measurements about the individual means, indicate the amount of day-to-day variation in the scores.

It has been shown (formula 24) that in terms of the estimated components of the population variance, the variance within individuals, V_{wI}, can be expressed as a sum of the daily component, σ_D^2, and the random component, σ_R^2:

$$V_{wI}=\sigma_D^2+\sigma_R^2$$

In the absence of significant differences between the daily means, i.e., when σ_D^2 is negligible, the V_{wI} can be used in place of the random variance, V_R.

Using V_{wI}, the coefficient of day-to-day consistency (r′) of repeated measurements may be expressed in the form of intraclass correlation:

$$r'=\frac{V_{bI}-V_{wI}}{V_{bI}+(k-1)\,V_{wI}} \qquad (37)$$

For details refer to reference No. 1.

BIBLIOGRAPHY

1. Alexander, H. W., Estimation of reliability when several trials are available, *Psychometrika* (in press).

2. Brozek, Josef, Harold Guetzkow, Olaf Mickelsen, and Ancel Keys, "Motor Performance of Normal Young Men Maintained on Restricted Intakes of Vitamin B Complex," *Journal of Applied Psychology*, 30:359-379, 1946.

3. Crump, S. L., "The Estimation of Variance Components in Analysis of Variance," *Biometrics Bulletin*, 2:7-11, 1946.

4. Fisher, R. A. *Statistical Methods for Research Workers.* Sixth edition. London: Oliver and Boyd, 1936. 339 pp.

5. Snedecor, G. W. *Statistical Methods.* Fourth edition. Ames, Iowa: Iowa State College Press, 1946. 485 pp.

Acknowledgment

This work was done in connection with research sponsored in part by the Sugar Research Foundation, New York, and the National Dairy Council, operating on behalf of the American Dairy Association.

Editorial Policies

THE Editorial Committee has approved the following policies for the publication of material in the *Quarterly:*

1. Manuscripts should be sent to the Editor who will see that each one is read by at least two members of the Board of Associate Editors. The Editor will advise the author as to the suitability of the paper or the desirability for revision.

2. Papers are not judged by arbitrary standards of length but on their content of new research results in the field of physical education, health education, and recreation, presented with the greatest brevity compatible with scientific accuracy and clarity.

3. Since manuscripts will not be insured against loss or damage, authors are expected to retain duplicate copies of all material submitted.

4. An original typewritten copy of the manuscript should be submitted. The content should be double spaced with a margin of 1½ inches on each side.

5. The author may include either a list of references at the end of the article or he may put them in footnotes or these two methods may be combined. Book publishers and periodicals do not always agree on the exact order of details in the preparation of references. Also, authors do not always include all the necessary information in references. For authors who have not published extensively a simplified form of magazine and book reference is shown. If a formal bibliography is included with the paper a simple footnote may be used as shown:[1]

1 Katherine B. Crisp, *Health for You,* p. 520.

In this case the footnotes are placed on each page and the bibliography is set up alphabetically and placed at the end of the article. The latter should show the facts of publication, i.e., edition if any, place published, publisher, date.[2,3]

2 Stern, Francis. *Applied Dietetics.* Baltimore, Md.: Williams and Wilkins Co., 1943. (form for book reference in a bibliography)

3 Corbin, H. D., "Current Problems in Recreation," *Journal of Health and Physical Education,* 15:6 (June, 1944). (form for magazine reference in a bibliography)

Note that the main items in a book reference are set off by periods while those in a magazine reference are set off by commas. These two types of references may be combined in what are called bibliographical footnotes which are placed on the corresponding pages as the article progresses. These are set up exactly as footnotes 2 and 3 except that page numbers are added at the end and, in this case, it is not necessary to put the author's last name first.

When there is more than one author to be listed in a formal bibliographical reference, the names are written with the Christian names or initials first in all cases except the first one when the last name appears first.[4]

4 Kraines, S. H., and E. S. Thetford. *Managing Your Mind.* New York: The Macmillan Co., 1944.

Footnotes should be separated from the text by lines running across the page. They should be inserted at the point in the page where the reference occurs and they should be numbered consecutively throughout the paper.

Although this form is preferred in the *Quarterly,* authors may submit articles with references prepared differently provided the essential information is given and the style used is that of well known journals.

There are many sources of information relative to the preparation of manuscripts for publication. A good source is *A Manual of Style* (10th Edition), Chicago, University of Chicago Press, 1927.

6. Charts, graphic tracings, etc., are required in many instances and best results are obtained by photographs of the size desired in reproduction and printed on glossy paper. Either originals or photographs may be submitted. Only those figures which aid in clarifying the article should be included. Authors are required to pay for cuts appearing in the *Quarterly* when the total is in excess of $25.00.

7. Tables and illustrations should be prepared with the size (4⅛x7 inches) of the *Quarterly* type page in mind, with the idea of conserving vertical space.

8. All illustrative material must be submitted in such form as to admit of photographic reproduction without retouching or redrawing. Plotted curves, guide lines, lettering, etc., should be done in india ink. Lettering should be large enough to be legible if illustrations have to be reduced in the process of making a cut.

9. Duplication of data in tables, charts, and protocols is not believed to be generally necessary. Too extensive use of such material is discouraged.

10. From time to time, the Editor may make requests for reviews and bibliographies. In this event, authors will be advised as to the amount of material desired.

Editorial Committee

Karl W. Bookwalter, David K. Brace, Frederick W. Cozens, Vivian Drenckhahn, C. H. McCloy, *Chairman*, Leonard A. Larson, Carl L. Nordly

Board of Associate Editors

Dorothy Ainsworth, V. S. Blanchard, David K. Brace, C. L. Brownell, Frederick W. Cozens, A. Lester Crapser, T. K. Cureton, Leslie W. Irwin, Peter V. Karpovich, C. V. Langton, William R. LaPorte, Leonard A. Larson, Frank S. Lloyd, C. H. McCloy, C. M. Miles, Gertrude E. Moulton, Jay B. Nash, D. Oberteuffer, W. W. Patty, Harry A. Scott, Jackson R. Sharman, George T. Stafford, Arthur H Steinhaus, Alden W. Thompson, W. W. Tuttle, Jessie F. Williams, Mary Wibel, Editor

REPRINTS FROM THE RESEARCH QUARTERLY

(1938 to date)

Reprints are available from the national office, 1201 16th St., N.W., Washington 6, D. C., for 10c each. Please order by number. All orders under $1.00 must be accompanied by funds.

1938

Q 203 Effect of Exercise, Fatigue and Exhaustion on the Electrical Potential of the Brain Cortex and Threshold of the Knee Jerk. Burge *et al.*

Q 204 Angles in Measuring Leg Lift. Carpenter.

Q 213 A Study of Falls in Skiing. Smart.

1939

Q 227 The C.C.C. as a Builder of Physical Fitness in Youth. Walker.

Q 229 Effect of the Elimination of the Center Jump. Frigard.

Q 232 Sociological Contributions of Physical Education to the Needs of the Negro. Bell.

Q 241 An Experiment in Self-Directed Study for College Freshmen. Guernsey.

1940

Q 246 Correlations of Measurements of Men Students. Ilsley.

Q 249 Prediction of Time in Swimming Based on Oxygen Consumption. Karpovich and LeMaistre.

Q 252 Prediction of the Advent of the Menarche from Height and Weight. Shuttleworth.

1941

Q 262 Intramural Athletics for Men in Negro Colleges. Taylor.

Q 267 Coefficient Performance Chart. Jefferson.

Q 270 Anthropometric Study of Negro and White Women. Steggerda and Petty.

Q 286 Race and Stature: A Study of Los Angeles School Children. Jones.

Q 288 The Factor Analysis as a Research Technique. McCloy.

Q 289 The Status of Student Health Programs in Negro Colleges. Cornely.

Q 293 Skeletal Symmetry in High Jumping. Krakower.

Q 306 A Fifty-nine Year Survey at Yale Reveals Freshmen are Becoming Younger, Heavier, and Taller. Deegan.

1943

Q 342 Determination of the Weight and Size of a Standard Discus for College Women. Tuttle.

Q 347 Yale University Completes One Year of Its Wartime Physical Training Program. Wickens and Murphy.

Q 352 Achievement Scale Scores for Wartime Swimming. Hewitt.

1946

Q 382 A Comparative Analysis of Secondary-School Boys' and Girls' Character and Personality Traits in Physical Education Classes. Blanchard.

Q 391 The Effect Upon Pulse Rate of Various Cadences in the Step-Up Test. Miller and Elbel.

Life and Honorary Members

American Association for Health, Physical Education, and Recreation

Abernathy, Ruth, State Education Department, Albany, New York.

Ainsworth, Dorothy S., 15 Barrett Place, Northampton, Massachusetts.

Anderson, Jackson M., Chief of Recreational Activities, American Airlines, Inc., 100 East 42nd Street, New York City 17.

*Anderson, W. G., Box 1911, Yale Station, New Haven, Connecticut.

Bascom, Frances R., Department of Physical Education for Women, University of Colorado, Boulder, Colorado.

Bell, Margaret, 15 Geddes Heights, Ann Arbor, Michigan.

Belshaw, R. E., Department of Physical Education, University of Washington, Seattle, Washington.

Blesh, T. Erwin, Payne-Whitney Gymnasium, Yale University, New Haven, Connecticut.

Bogatko, Paul P., 23 Clendenny Avenue, Jersey City 4, New Jersey.

Bookwalter, Carolyn W., 527 Highland, Bloomington, Indiana.

Bookwalter, Karl W., Director, Service and Research, School of Health, Physical Education, and Recreation, Indiana University, Bloomington, Indiana.

Booth, Minnie Anne, 135 Navarre Road, Rochester, New York.

Bourgeois, Mrs. Vesta Richard, Box 207, Lafayette, Louisiana.

Cole, Eleanor M., Stockton Junior College, Stockton, California.

Cureton, Thomas K., Jr., University of Illinois, Urbana, Illinois.

*Day, William E., 1377 Princeton Avenue, Salt Lake City, Utah.

Deckman, Beatrice F., 32 Cooper Street, Westmont, New Jersey.

Delta Theta Epsilon, University of Illinois, Urbana, Illinois.

Di Filippo, Victor J., Department of Health and Physical Education, Seton Hall College, South Orange, New Jersey.

Di Giovanna, Lt. Vincent, PT & ADQMS, Staff and Faculty, Camp Lee, Virginia.

Di Napoli, Frank, 880 Beardsley, Akron 11, Ohio.

Dodd, Gilbert B., Director of Physical Education, Taylor University, Upland, Indiana.

Drew, A. Gwendolyn, Physical Education for Women, Washington University, St. Louis 5, Missouri.

Duncan, Norman D., University of California, Los Angeles, California.

DuPain, George, Rose Bank, 158 Parramatta Road, Ashfield, N.S.W., Australia.

Durbin, Margaret E., Nazareth College, Nazareth, Kentucky.

Elliott, E. S., Columbia University, New York City.

Espenschade, Anna, Hearst Gymnasium, University of California, Berkeley, California.

Fontaine, Carnie B., Canby Elementary Schools, Canby, Oregon.

Francis, Robert J., University of Wisconsin, Madison.

Frankel, Mrs. Elkan F., 86-38 162nd Street, Jamaica 2, New York.

French, Esther, Illinois State Normal University, Normal, Illinois.

Friermood, Harold T., Y.M.C.A., 347 Madison Avenue, New York City.

Fulton, Ruth E., University of California, Los Angeles, California.

Gamboa, Jesus Ferrer, University of Panama, Apartado Postal num, 738, Panama, R. de P.

Garrison, Jessie R., State Department of Education, Montgomery, Alabama.

Gates, Grover, 835 Hampton Way, Fresno, California.

Giauque, Comdr. C., 310 West Montgomery Avenue, Rockville, Missouri.

Golomb, Jacob J., 26 East 14th Street, New York City 3.

Greene, Margaret D., 2077 North Beverly Glen, Los Angeles, California.

Gugisberg, Mercedes, University of New Mexico, Albuquerque, New Mexico.

Hansma, Jack, Western Maryland College, Westminster, Maryland.

Hepp, Frank, Royal Hungarian College of Physical Education, Gyori ut 13, Budapest, Hungary.

Hester, Mrs. Clara L., 3953 Graceland Avenue, Indianapolis, Indiana.

Hill, Laurence S., Ithaca College, Ithaca, New York.

Houston, Ruth, State Teachers College, Buffalo, New York.

*Honorary members.

Howe, Arthur L., Hamburg, New York.
Ierardi, Thomas, City College, New York City.

Irwin, Leslie W., Department of Physical Education, 84 Exeter Street, Boston University, Boston 16, Massachusetts.

Johnson, Louis R. W., 512 North Alfred Street, Alexandria, Virginia.

Kasch, Fred W., 715 South Wood Street, Chicago, Illinois.

Kirsten, Arnold F., Jewish Community Center, 45 Market Street, Lynn, Massachusetts.

Kranz, Leon G., Northwestern University, Evanston, Illinois.

Lagarde, Rene, 3190 Van Horne Street, Apartment 7, Montreal 26, P.Q., Canada.

Langton, Clair V. Oregon State Agricultural College, Corvallis, Oregon.

Lapp, V. W., 2234 Cotner Boulevard, Lincoln, Nebraska.

La Tourrette, Charlotte, Route No. 4, Athens, Ohio.

Lee, Mabel, University of Nebraska, Lincoln, Nebraska.

Lewis, William F., Master of the Sword, United States Military Academy, West Point, New York.

Luehring, F. W., 435 Riverview Road, Swarthmore, Pennsylvania.

Mac Gregor, John M., 309 South Market St., Johnstown, N. Y.

Mackey, Anne D., 38-30 211 Street, Bayside, Long Island, New York.

Manley, Helen, U. S. Office of Education, Federal Security Agency, Washington 25, D. C.

Marting, Catherine L., Pomerene Hall, Ohio State University, Columbus, Ohio.

Matson, Elizabeth, College of Pacific, Stockton, California.

McCloy, C. H., State University of Iowa, Iowa City, Iowa.

*McCurdy, Mrs. Persis B., 1235 Brickell Ave., Miami, Fla.

McKinstry, Helen, Russell Sage College, Troy, New York.

McLeod, Wilma, Box 396, Ontario, Oregon.

McNeely, Simon A., Department of Education, Baton Rouge 4, Louisiana.

Metcalf, T. N., 5724 Kenwood Avenue, Chicago, Illinois.

Miller, Ben W., 1201 16th Street, Washington 6, D. C.

Miller, June M., 312 12th Street, Menomonie, Wisconsin.

Mitchell, E. D., University of Michigan, Ann Arbor, Michigan.

Mohr, Dorothy R., Texas State College for Women, Denton, Texas.

Moore, Rodgers, Box 823, Porterville, California.

Moorhead, W. G., 121 Center Street, East Stroudsburg, Pennsylvania.

Moulton, Gertrude E., 291 Forest Avenue, Oberlin, Ohio.

Nason, Raymond E., 520 Kenbrook Road, Worthington, Ohio.

Nass, Paul V., 11976 San Vicente, Los Angeles 24, California.

Neilson, N. P., University of Utah, Salt Lake City, Utah.

Norris, J. Anna, 1429 East River Road, Minneapolis, Minnesota.

Noyes, Elizabeth, R.F.D. 3, Box 56, Putnam, Connecticut.

Olds, Lloyd W., Hurron Apartments, 921 Westcross Street, Ypsilanti, Michigan.

Osborn, Lola L., 290 East Washington Avenue, Chico, California.

Panzer College of Physical Education and Hygiene, 139 Glenwood Avenue, East Orange, New Jersey.

Pille, Roy F., Miami University, Oxford, Ohio.

*Posse, Baroness Rose, Shelton Hotel, Lexington Avenue at 48th Street, New York City.

Priest, Ernest G., 907 North Reus Street, Pensacola, Florida.

Rafeld, Jackson W., 2527 South Freedom Avenue, Alliance, Ohio.

Ralston, B. A., Dalhousie University, Halifax, Nova Scotia, Canada.

Rearick, Elizabeth C., MacMurray College for Women, Jacksonville, Illinois.

Rector, Ruth V., Central High School, Xenia, Ohio.

Rominiecki, Claudia Ann, 450 Camden Avenue Moorestown, New Jersey.

Ross, Brenda B., 432B Veteran Avenue, Los Angeles 24, California.

Rugen, Mabel, University of Michigan, Ann Arbor, Michigan.

Ruoff, Daniel H., Ada, Minnesota.

Schnell, H. W., College of Physical Education, Health, and Athletics, University of Florida, Gainesville.

Schuyler, Gretchen, Sargent College, 6 Everett Street, Cambridge 38, Massachusetts.

Scott, M. Gladys, State University of Iowa, Iowa City, Iowa.

Sellers, Dorothy G., Winthrop Apt. No. 8, 2332 10th Court, South, Birmingham, Ala.

Shaw, John H., Department of Physical Education, Syracuse University, Syracuse 10, New York.

Siller, J. Granville, 589 San Luis Road, Berkeley 7, California.

Silver, Joseph T., 1387 Harding Terrace, Hillside, New Jersey.

Sinclair, Caroline B., College of William and Mary, Williamsburg, Virginia.

Snyder, Raymond, Director of Health and Physical Education, Willimantic State Teachers College, Willimantic, Connecticut.

Sparks, Lestle J., 1045 North Fourteenth Street, Salem, Oregon.

Stafford, Grace M., N.Y.A., 611 Arlington Place, Chicago, Illinois.

*Stecher, William A., c/o C. F. Weeber, Box 3468, Honolulu, T. H.

Steinhaus, Arthur, Robinson Foundation, Inc., New York City.

Streit, W. K., Board of Education, Cincinnati, Ohio.

Swain, Leslie E., c/o John Hay Library, Brown University, Providence, Rhode Island.

Tarr, Edna V., Pacific University, Forest Grove, Oregon.

Titus, Mary, 629½ Eleventh Avenue, Huntington, West Virginia.

Torregrosa, Felicio M., University of Puerto Rico, Rio Piedras, Puerto Rico.

Turner, Clair E., 19 Village Lane, Arlington, Massachusetts.

Walker, Charles L., State College, San Jose, California.

Williams, Lafayette, 226 Olive Street, Edwardsville, Illinois.

Wood, Calvin, Caesar Rodney High School, Camden, Delaware.

Young, Olive G., University of South Dakota, Vermillion, South Dakota.

AMERICAN ASSOCIATION FOR HEALTH, PHYSICAL EDUCATION, AND RECREATION

(A Department of the National Education Association)

BOARD OF DIRECTORS

President, Vaughn S. Blanchard, 467 West Hancock Avenue, Detroit, Michigan.
President-Elect, Ruth Evans, Public School System, Springfield, Massachusetts.
Past President, Helen Manley, Public Schools, University City, Missouri.
Vice President, Health Education, Arthur H Steinhaus, George Williams College, Chicago 15, Illinois.
Vice President, Physical Education, Louis Keller, University of Minnesota, Minneapolis.
Vice President, Recreation, Caswell Miles, State Education Department, Albany 1, New York.
Central District: Carl L. Nordly.
Eastern District: George W. Ayars.
Midwest District: Iris Boulton.
Northwest District: Clair V. Langton.
Southern District: Thomas E. McDonough.
Southwest District: Eleanor Metheny.

ADDITIONAL MEMBERS OF REPRESENTATIVE ASSEMBLY

Sections of Health Education Division
Dental Health: No report.
Safety Education: No report.
Health Instruction: No report.
Mental Health: Vincent DiGiovanna.
School Nursing: No report.
Nutrition Education: Alice Smith.
School Physicians: Louis R. Burnett.
Therapeutics: Ellen Kelly.

Sections of Physical Education Division

College and University Physical Education for Men: John Shaw.
College and University Physical Education for Women: Mary McKee.
Dance: No report.
Intramural Athletics: John Van Why.
Men's Athletics: H. H. House.
Private School Physical Education: No report.
Public School Physical Education: No report.
Women's Athletics: Martha Gable.

Sections of Recreation Division

Camping and Outing: No report.
College Recreation: John Scannell.
Industrial Recreation: No report.
Institutional Recreation: No report.
Municipal Recreation: No report.
Park Recreation: No report.
Private Recreation: Harry D. Edgren.
School Recreation: Ruth Abernathy.

General Sections

City Directors: Pattric Ruth O'Keefe.
Measurement and Evaluation: Esther French.
Professional and Public Relations: Howard Danford
Professional Education: Ralph Piper.
Research: Thomas K. Cureton.
Student: Nancy Neville.

Unofficial

Aquatic: Jack Tourney.
Elementary Physical Education: Edwina Jones.

District Association Presidents

Central: Elizabeth Graybeal.
Eastern: Clifford L. Brownell.
Midwest: Leon G. Kranz.
Northwest: Dorothea Lensch.
Southern: Joy W. Kistler.
Southwest: Verne Landreth.

State Representatives

(Central District)

Colorado (2): Maisie Hoyt; Willard N. Greim.
Iowa (2): Theresa Anderson; Roy Moore.
Kansas (2): L. P. Washburn; Edna McCullough.
Minnesota (2): P. J. Sandell; Mabel J. Shirley.
Missouri (3): Muriel M. Lomax; Helen Fahey; Donald Ingerson.
Nebraska (1): Ed Higginbotham.
North Dakota (1): E. J. Cassell.
South Dakota (1): Irv Nelson.
Wyoming (0): None.

(Eastern District)

Connecticut (2): Joseph V. Burns; Josephine Rogers.
Delaware (1): George W. Ayars.
District of Columbia (2): Julian N. Colangelo; Virginia Dennis.
Maine (1): Howard G. Richardson.
Maryland (1): Thomas C. Ferguson.